COORDINATED TRANSLATION PROGRAMS REDUCE PROLIFERATION AND PROMOTE THE EPITHELIAL TO MESENCHYMAL TRANSITION (EMT)

A Dissertation

Presented to the Faculty of the Weill Cornell Graduate School

of Medical Sciences

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Randall Andrew Dass

January 2018

COORDINATED TRANSLATIONAL PROGRAMS REDUCE PROLIFERATION AND PROMOTE THE EPITHELIAL TO MESENCHYMAL TRANSITION (EMT)

Randall Andrew Dass, PhD

Cornell University, 2018

ABSTRACT

It is currently thought that uncontrolled cellular proliferation as a result of cellular insults be they genetic or environmental is the major driving force behind the development of malignant neoplastic growth and subsequent metastatic dissemination. Based on the current paradigm and thinking therapeutic interventions for cancers specifically target rapidly dividing cells. Here, within this body of work is described the role that rDNA transcription plays in both the proliferative aspect of cancer progression as well as the role it plays in the metastatic dissemination of cancerous cells. Described herein is the first global translational analysis of the EMT program and the description of the coordination of two endogenous translation programs; one program that drives proliferation and another, independent program that drives differentiation by the translation of pro-mesenchymal and pro-migratory genes.

BIOGRAPHICAL SKETCH

Randall A. Dass began his interest in science and scientific discourse under the patient and enthusiastic tutelage of Raphael Ramlal. He strongly encouraged looking at problems from different perspectives and was always willing to answer questions regardless of their validity. Randall's interest in science and specifically biology was encouraged and strengthened while in attendance at The College of the Immaculate Conception (CIC) in Port of Spain, Trinidad, West Indies.

Randall continued on the path of scientific interest by earning a Baccalaureate in Science (BSc.) with honours in Molecular Biology and Genetics at the University of Guelph in Ontario, Canada. This was followed by graduate research with the Department of Animal and Poultry Science at Guelph. Upon returning to the United States of America, Randall was employed as a technician in the laboratory of Dr. Anthony Brown, where he was exposed to cancer biology, cellular signaling, ribosome biogenesis and the epithelial to mesenchymal transition.

Having been strongly interested in translation as to how it pertains to cancer progression, Randall joined the laboratory of Dr. Scott Blanchard to pursue a PhD in cancer biology and translation. He was able to implement a modified ribosome profiling protocol in order to get the first global look into how translation is altered as cells change their identity from an epithelial state to a mesenchymal state.

To Aria, without your love and support this journey would not have been possible. Your patience, understanding and your proofreading skills were indispensable. Thank you for listening to and paying attention to all my practice presentations.

To my family, thank you for all your support and the opportunity to undertake a journey such as this one.

Grandma, I hope you can see this achievement and it makes you proud.

ACKNOWLEDGEMENTS

I would like to acknowledge the support and mentorship that was offered to me by Dr. Scott Blanchard. His attention to detail and ability to convey scientific ideas and scientific rigor, all skills he sought to improve in me, have made me a much more capable scientist. I'll be forever grateful for the sense a capability and confidence he has instilled me with. I would also like to acknowledge Dr. Blanchard for creating an extremely focused and productive laboratory environment.

I would also like to acknowledge the mentorship of Dr. Theresa Vincent who encouraged me to read scientific literature, discuss scientific ideas and eventually encouraged me to pursue a career in science. Much of my knowledge in cancer biology and EMT I have gained from her.

I would like to acknowledge the members of my committee, Dr. Andrea Ventura for his insights into RNA biology and especially Dr. David Lyden for his experimental advice, his tutelage on cancer biology and metastasis and perhaps most importantly for his concern for my wellbeing and my future development.

I would like to acknowledge Roger Altman for taking care of all of the tasks that allows all the members of the lab to focus mainly on science. I would also like to acknowledge all members of the Blanchard lab, past and present, especially Dr. Benjamin J. Burnette and Dr. Jose Alejo Amaya.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

EMT	Epithelial-to-Mesenchymal Transition
MET	Mesenchymal-to-Epithelial Transition
CSC	Cancer Stem Cells
RNA	ribonucleic acid
DNA	deoxyribonucleic acid
mRNA	messenger ribonucleic acid
tRNA	transfer ribonucleic acid
rDNA	ribosomal deoxyribonucleic acid
rRNA	ribosomal ribonucleic acid
ssDNA	single stranded deoxyribonucleic acid
ssRNA	single stranded ribonucleic acid
TGFβ	Tissue Growth Factor Beta
NMuMG	Namru Mus musculus Mammary Gland
PyMT	polyoma middle T antigen
NOR	Nucleolar Organizing Region
NGS	Next Generation Sequencing
RNASeq	Ribonucleic acid sequencing
PCR	Polymerase Chain Reaction
FISH	Fluorescence In Situ Hybridization
FRET	Fluorescence Resonance Energy Transfer
smFRET	single molecule Fluorescence Resonance Energy Transfer
Cryo-EM	cryogenic electron microscopy
Pol I	RNA Polymerase I
Pol II	RNA Polymerase II
Pol III	RNA Polymerase III
UBF	Upstream Binding Factor
pUBF	phosphorylated Upstream Binding Factor
Snai1	Snail1
Cdh1	E-cadherin
Cdh2	N-cadherin
Vim	Vimentin

CK8	Cytokeratin 8
Fib	Fibrillarin
snoRNAs	small nucleolar RNA
miRNA	micro RNA
lncRNA	long non-coding RNA
FUrd	5’-fluorouridine
EdU	5-Ethynyl-2´-deoxyuridine
EU	5-Ethynyl-2´-oxyuridine
DAPI	4',6-diamidino-2-phenylindole
Eif6	Eukaryotic initiation factor 6
Eif4a1	Eukaryotic initiation factor isoform a1
Eif4a2	Eukaryotic initiation factor isoform a2
Eef2	Eukaryotic elongation factor 2
Eef1b2	Eukaryotic elongation factor beta 2
Eef1d	Eukaryotic elongation factor gamma
Eef1g	Eukaryotic elongation factor delta
Eef1a	Eukaryotic elongation factor isoform 1a

General Introduction

Cancer is a worldwide problem that has been increasing in incidence as the worldwide population grows and ages[1]. There has been a significant amount of energy, research and time spent in trying to decipher the complexities of the many facets of cancer[2]. Progress has been made in understanding how tumors may be formed and what biological alterations can drive carcinogenesis[1]. The lines of evidence have pointed toward treatments that aim at stopping the primary tumor from growing and have mainly steered the research premises about cancer in the direction that the only problem is one of proliferation.

There has slowly been a paradigm shift as more research that perhaps one of the most damaging aspects of cancer progression is the development of metastases where in some cancers like breast cancer 90% of the fatalities can be attributed to metastatic spread[3]. Building on these observations more research has been focused on delineating the molecular underpinnings of how the metastatic cascade develops, how cells survive in circulation, home into and colonize distant sites and finally how they begin to form secondary tumors with in the newly compromised tissue[4–6].

Cancer has been thought of as both a simple and a complex disease, simple in that a virus can cause tumorigenesis and complex in that many biological and cellular check points and safe guards must be overcome in order to for cancer development and progression to occur[7–10]. In this respect cancer can be very complicated with multiple cellular insults giving rise to tumors and same principles may apply to the metastatic cascade. What has been observed however in both instances is that normal cellular systems can be co-opted in order to promote both proliferation and metastasis.

Though complex in nature and most likely the result of multiple programs acting in a synergistic manner we have chosen to focus on the Epithelial to Mesenchymal Transition (EMT) as potentially one of the driving forces that could facilitate cells with the necessary phenotypic characteristics needed in order for cells to leave the primary tumor[11]. The process of EMT involves cells losing their junctional adhesion molecules, rearranging their cytoskeletal structure, becoming more invasive and adopting the morphology of mesenchymal cells[12,13].

The process of EMT was first identified in the early 1960's as a naturally occurring developmental process[12–14]. During normal developmental processes the EMT program is executed in order to allow for the correct migration of cells into the correct areas for specific tissue development. It is however thought that the re-activation of the EMT program during adulthood leads to disease states[15,16]. As the process of EMT has been studied and the molecular mechanisms identified it has mostly been studied as a transcriptionally driven process[17–19].

There have been several signaling pathways that have been identified and studied that drive or initiate the EMT program, they are Wnt, TGFβ and Hypoxia (Notch)[20–23]. Within the developing chick neural tube the EMT is considered a Wnt driven system whereas there is evidence of TGFβ and Hypoxia driven EMT in mammary epithelia, particularly in the context of a disease state[20,24–26]. Despite having different initiation signals there are a number of key events that typify an EMT program being activated of note and perhaps of key functional consequence is the loss of junctional proteins such as E-cadherin and β-catenin and the concomitant gain of mesenchymal proteins such as Vimentin and N-cadherin[17,27–29].

There has been prior work investigating possible translation control of specific proteins during the EMT process but there has yet to be any global translational analyses[30,31]. This was most likely owing to the lack of a suitable technique that could aid in the global analysis of which genes are actively translated as well as how genes are translated. In 2009 the ribosome profiling technique was introduced which allows for a global look at not only which genes are being actively translated but how they are translated[32–35].

Given the extensive nature of the transcriptional reprogramming that occurs as the EMT program is executed it must be complemented by a large scale translational change as a cell changes identity from an epithelial cell to a mesenchymal cell. In order to build the repertoire of mesenchymal proteins needed for the cellular identity switch to occur it would be necessary for there to be a coordinated translational program to facilitate the making of the requisite proteins. In light of this hypothesis, that large scale translation changes, occur during EMT we sought to use a modified ribosome profiling method in order to get the first global snapshot of translation changes during the execution of the EMT program.

CHAPTER 1

IMPLEMENTATION OF THE RIBOSOME PROFILING METHOD

1.1 OVERVIEW

In attempting to address the problem of metastatic spread, we endeavored to better understand the epithelial to mesenchymal transition (EMT). Though metastasis is a complex process, it is thought that EMT is one of the first processes to occur early on in tumour development that could lead to the dissemination of cancerous cells.

In attempting to further delineate the molecular mechanisms of the EMT process we made a paradoxical observation, that rDNA transcription is increased despite a lack of proliferation. This is, to the best of our knowledge the first instance of rDNA transcription not corresponding with a cellular proliferation program. This finding is extensively discussed in the manuscript Prakash *et al.* of which I am a co-author.

Given that new rDNA transcription gives rise to new rRNA and by extension new ribosomes we hypothesized that there must be global translation changes that would accompany a change in the cellular identity of one cell from an epithelial cell to a mesenchymal cell. In order to get a global snapshot of which genes are being translated as cells execute the EMT program we implemented the ribosome profiling protocol.

Ribosome profiling was first described in a landmark paper by Ingolia et al in 2009 and offered a primary look into the power and scope of the new technique. By leveraging the power of next generation sequencing (NGS), the authors were able to come up with a strategy that minimizes bias caused by the

ligation protocols of most NGS systems, particularly the protocol used by the Illumina company for RNASeq library preparation[32].

The minimization of ligation bias was not the predominant improvement of this protocol however, although the changes developed in the protocol allowed for the overcoming of a large technical challenge, the true power lies in the development of a technique that can measure translation of mRNA transcripts. This is novel and it adds another level of biological information when combined with doing RNASeq studies. When combined with the transcription background this technique can show not only *which* transcripts are actively being translated but *how* they are translated. In addition to giving an idea of which transcripts are being translated that it can give information on upstream open reading frames (uORFs), alternative start sites for translation, truncated peptides, alternative reading frames and novel proteins[35–38]. This protocol was formalized and improved and published in 2012, it is this improved protocol that we have chosen to follow and modify to operate in the lab[33].

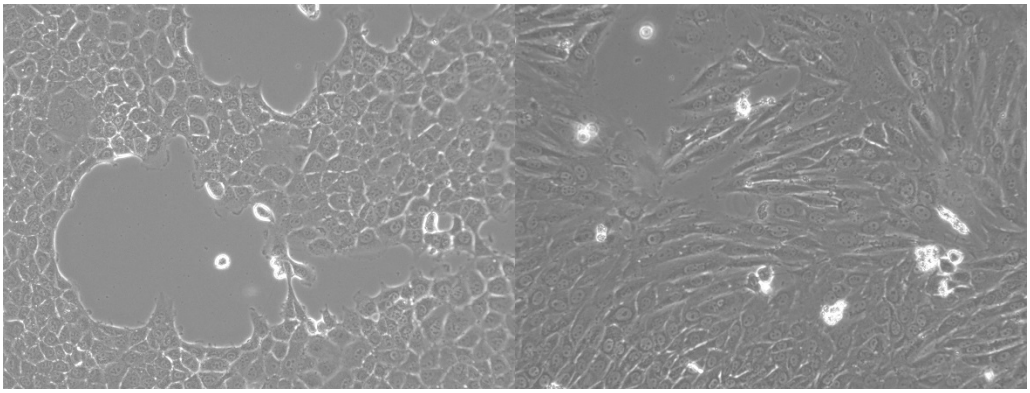

Figure 1. Morphology of vehicle treated (left panel) or TGFβ treated (800 pM) NMuMG cells (right panel) as seen at 20X magnification using phase contrast objective.

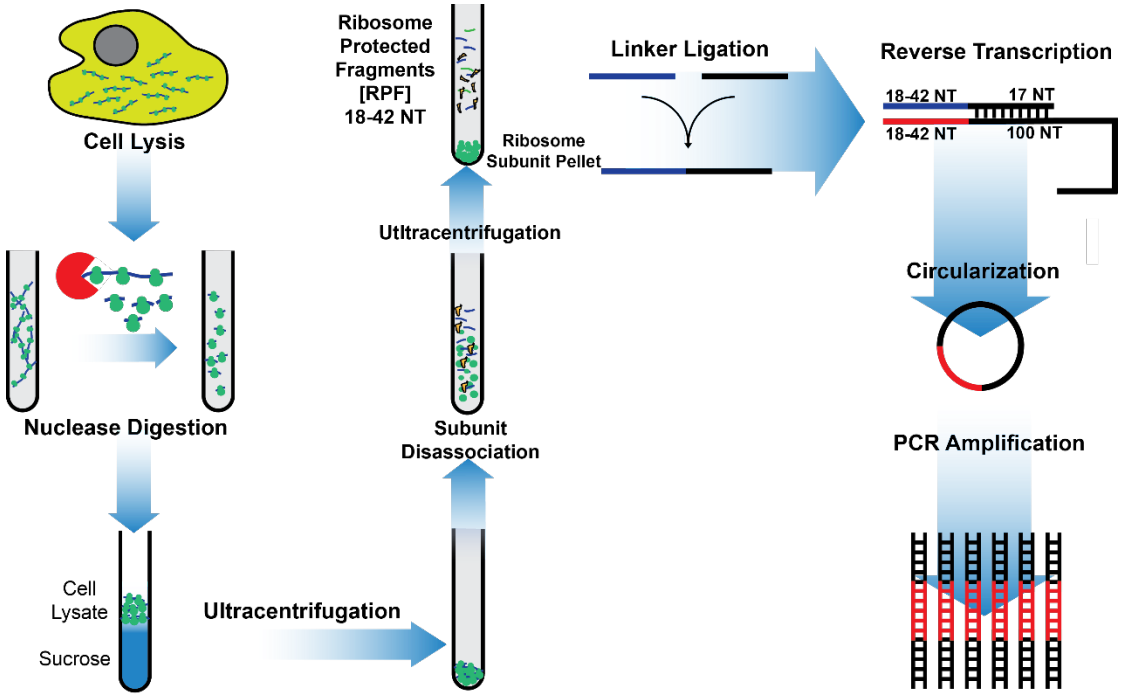

Figure. 2. Schematic diagram of the work flow involved in generating Illumina compatible libraries from cell derived ribosome protected fragments. There are two main modifications to the protocol, alteration of centrifugation steps by the addition of a sub unit separation step and subsequent second ultra-centrifugation.

In the following protocol the NMuMG EMT cell system was extensively used as we were trying to determine what translational control, if any, takes place as cells go through EMT.

1.2 METHOD

1.2.1 Cell Growth and Harvesting:

NMuMG cells grown for ribosome profiling are grown in one (1) T225 flask per biological repeat.

Three (3) vehicle treated repeats are grown in three (3) T225 flasks. The same would be done for TGFB1 treated cells.

From one (1) T225 flask at 80 – 90% confluence (confluency here refers to what percentage (on average) of the 225 square cm^2 area is covered by cells) would yield approximately **40 million cells**.

Treat the cells with cycloheximide (CHX) @ 350μM final concentration for 5 mins @ 37 degrees (in the incubator).

Trypsinize the cells with trypsin/EDTA (Cat#25300054) to lift the cells off of the surface of the flask. The NMuMG cells are adherent, thus they will stick to the surface).

Gently (< 200 x g) pellet the cells at four degrees Celsius.

Re-suspend the cells in sterile PBS (pH 7.4) in order to wash off any remaining trypsin or media. Pellet again gently (<200 x g).

Re-suspend the cells in 8mL of sterile PBS (pH 7.4). I place 1 mL into a clean, labelled, sterile Eppendorf tube.

Pellet the cells in a table top picofuge, aspirate as much of the PBS off as possible then **flash freeze** in liquid nitrogen.

Store at **-80°C indefinitely**.

On average, each aliquot will contain approximately **5 million cells**. Using this methodology usually have 8 aliquots **PER BIOLOGICAL REPEAT**.

As far as the protocol goes I will outline the Ingolia et al protocol 2012 and emphasize what is unique to our lab, what I have changed and what is the same as the protocol. Unless otherwise stated, the composition of the buffers has not been altered and should be made and used as described in the protocol.

1.2.2 Cell Lysis:
(Timing: 1 hour)

Re-suspend frozen cell pellet in 400 µL of lysis buffer.

Lysis buffer:

20 mM Tris (pH 7.4)

150 mM NaCl

5 mM MgCl

1 mM DTT

350 µM Cycloheximide

1% Triton X-100 (vol/vol)

Turbo DNase 25 U mL^{-1}

Thus far, I have only attempted the lysis without flash freezing as I have only used a mammalian cell line and have had no problems with cell lysis. Additionally, another graduate student (Leyi Wang) in our lab has used another lysis buffer mixture that gave similar yields of ribosome pellet to the one in the protocol (NP-40 0.5% final concentration and sodium deoxycholate 0.5% final concentration) of ribosomal pellets. With that in mind, I have chosen to use the lysis buffer from the protocol.

1. After placing the cell lysate into a clean non-stick 1.5 mL micro-centrifuge tube pipette cells up and down in lysis buffer to re-suspend and disrupt cell clumps, the more surface area the lysis buffer has to act the better the lysis reaction will occur. Incubate cells in the lysis buffer on ice for 10 mins

2. Triturate cells by passing them through a 25 7/8th gauge needle 10 times.

This serves to further break down the cell membranes and to physically shear genomic DNA (gDNA). This becomes important later on as the gDNA is degraded by DNAse treatment. As physical shearing breaks down the gDNA creating smaller fragments and a larger surface area foo the DNAse to work on.

3. Clarify lysate by pelleting cell debris by centrifugation for 10 mins @ 20 000g @ 4 °C. The Eppendorf 5424R and at 15 000 rpm gets to 20 000g.

1.2.3 Development of an enzymatic activity assay for Rnase I:

As stated in the 2012 protocol[33] published in Nature protocols by Ingolia et al. one of the most sensitive aspects of the protocol is the generation of the ribosome protected footprint (RPF) using the enzyme Rnase I. In order to be able to detect and eliminate any lot to lot variation or to be able to test the enzymatic activity of the nuclease it was necessary to develop an activity assay where the activity of enzyme could be quantified.

It was hypothesized that the ssRNA size markers when bound to SYBR Gold would have a measurable signal in the 1µM concentration regime. This hypothesis could be directly tested using the visible blue light transilluminator used to take images of the RNA gels.

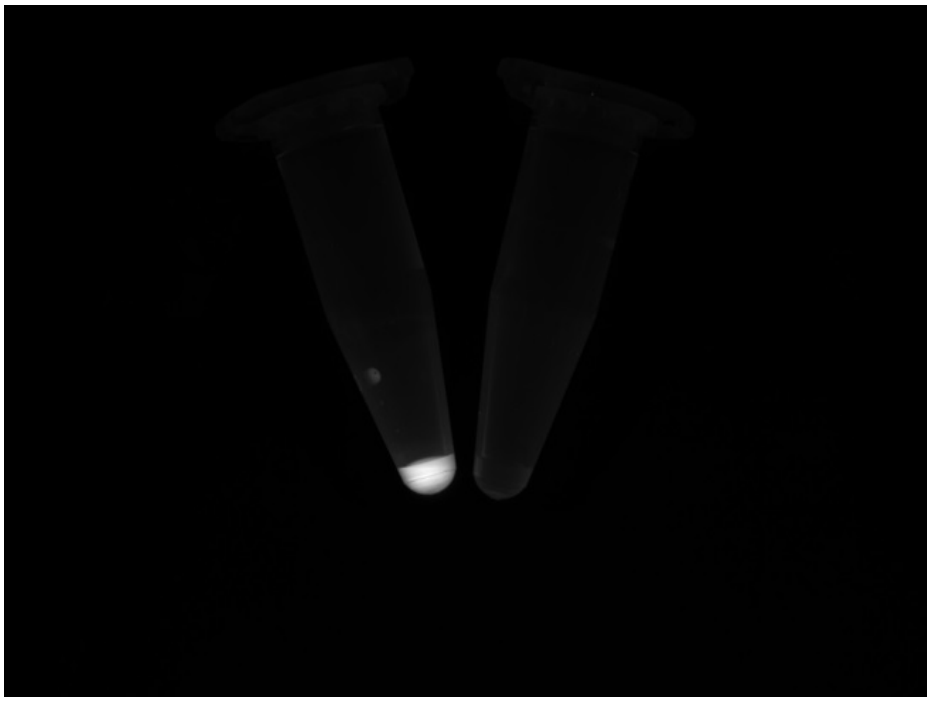

Figure 3. Left Eppendorf tube contains RNA (1µM) and SYBR Gold which fluoresces upon exposure to visible blue light whereas the right tube contains SYBR gold only and is devoid of any noticeable auto-fluorescence.

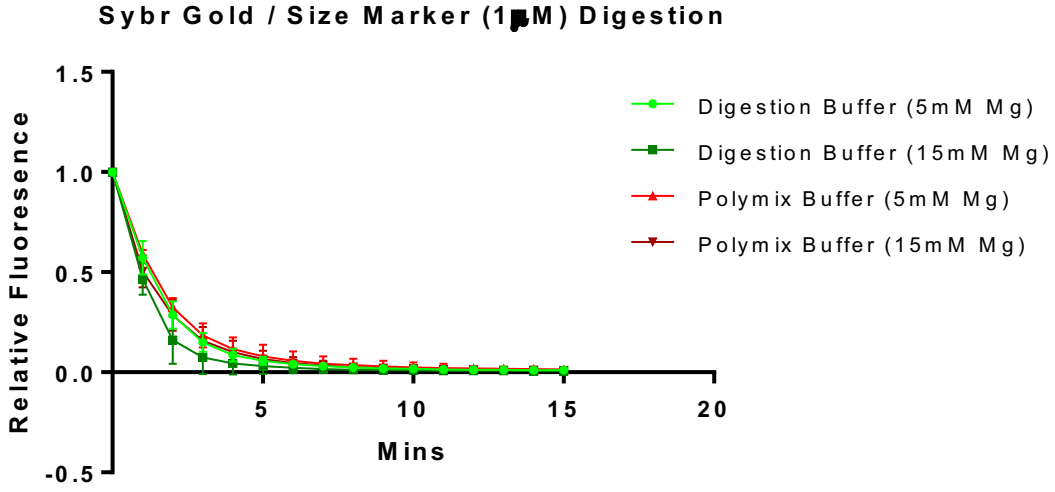

Figure 4. Quantification of Rnase I activity by loss of fluorescence over time in different buffer conditions using a Tecan Infinite M1000 Pro microplate reader.

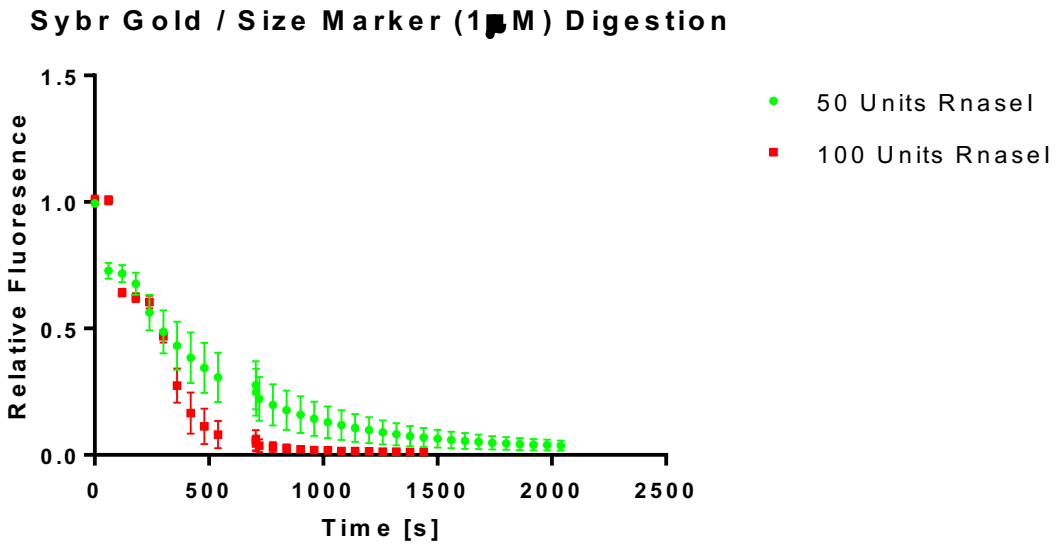

Figure 5. Quantification of Rnase I activity by loss of fluorescence over time with differing amounts of nuclease using a Tecan Infinite M1000 Pro microplate reader.

1.2.4 Nuclease Foot-printing:

(Timing: 9 hours)

4. Transfer 300 µL of clarified lysate to a new, clean non-stick 1.5 mL micro-centrifuge tube. To the 300 µL add 7.5 µL of RNase I (100 U μL^{-1}). Incubate with gentle mixing @ room temperature for 15 mins. BE CAREFUL with tips, **THEY ARE CONTAMINATED with RNase I**, treat as if they are radioactive. Discard tips into a 50mL Falcon tube containing 0.01%SDS/10 mM acetic acid solution. This tube is to be thrown out at the end of the digestion period. In order to use either the nutator or the rotator I wipe the outside of the tubes down with 0.01% SDS/ 1 mM acetic acid and I use the rotator in cell culture.

5. Add 10 µL of SUPERase.In RNase inhibitor to stop the nuclease digestion. Pipette up and down several times to mix. BE CAREFUL with tips, **THEY ARE CONTAMINATED**, treat as if they are radioactive. Discard tips into a beaker containing 0.01%SDS/10 mM acetic acid solution.

ALTERATION: Instead of using the polysome buffer to make the 1M sucrose cushion we have opted to use a high salt buffer that consists of:

1M NH_4Cl

20 mM Mg^{2+}

500 µM Cycloheximide

The rationale behind this is that we can both inhibit the RNase I enzyme with the high salt. In addition, the higher cycloheximide and magnesium concentrations aids with keeping the ribosome locked to the message transcript mRNA.

6. Transfer the digestion into a 13 X 51 mm polycarbonate ultracentrifuge tube. **Underlay** 0.9 mL of the 1M sucrose cushion under the sample. The sample lysate will float on top of the sucrose cushion, there will be a distinct and visible interface between the two layers. Mark the outside of the tube where you expect the ribosome pellet to appear.

7. Pellet the ribosomes by centrifugation in a TLA100.3 rotor @ 78 000 r.p.m. @ 4 °C for 4 hours. The TLA100.3 rotor is kept in the cold room on the bench to the right as you enter the door.

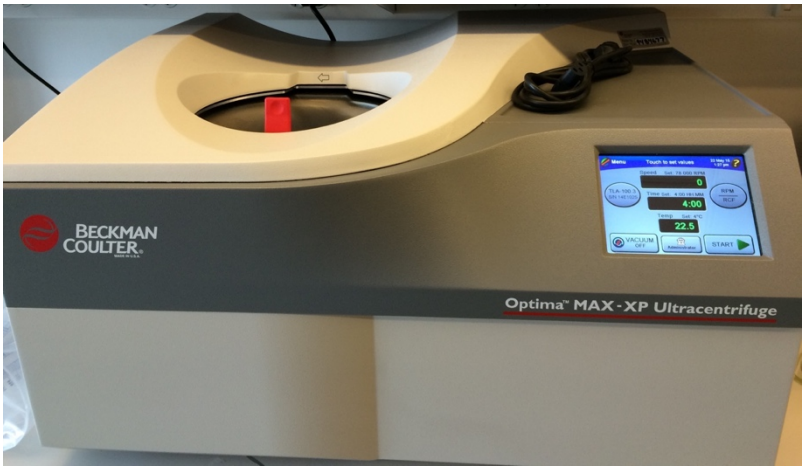

Figure 6. Picture of the benchtop ultra-centrifuge used to pellet nuclease generated monosomes or pelleting of ribosome sub-units.

Figure 7. Picture of ribosome pellet (just above black line) in 51 x 13 mm polycarbonate ultracentrifuge tube.

ALTERATION:

After seeing the published gels and looking at the gels I produced, it occurred to us that we can use a dissociation buffer to break apart the ribosome to release the mRNA fragments. This would give us gels with less background noise as it eliminates high molecular weight (HMW) RNA and it may eliminate the ribosomal RNA (rRNA) fragments that form a large part of the contaminating pool of RNA fragments that have to be depleted later on. In addition to that it eliminates the need for TRIzol purification of the samples and it cuts down on RNA processing time.

2.2.5 Sub-unit Dissociation:

0.5 M KCl

1mM Puromycin

SUPERase.In 100U / mL of buffer.

The puromycin aids in peptide release[39]. The disadvantage of this alteration is that without the TRIzol there is no protein denaturation, this means that there may be active RNase I particles that may have come down with the ribosome pellet and they would then be released with the dissociation buffer and the be free to degrade the RPFs. In order to overcome this I do a subsequent wash of the pellet with dissociation buffer before physically disrupting the ribosome. Measure the diameter of the ribosome pellet with a ruler or Vernier caliper.

8. Add 300 µL of dissociation buffer to the ribosome pellet. Incubate at room temperature pipette up and down until you see the ribosome pellet break up. This may take up to a minute of constant pipetting.

9. Centrifuge again in the ultracentrifuge @ 90 000 r.p.m. @ 4 °C for 2 hours.

Figure 8. Picture of ribosome sub-unit pellet (right above black line) in 51 x 13 mm polycarbonate ultracentrifuge tube. Note the slight brown discoloration in the pelleted subunits. This is typical of the pelleted sub-unit pellets.

10. Carefully remove the supernatant, avoid the ribosome pellet. The supernatant contains your RPFs. In order to purify and concentrate the RNA the subsequent steps describe the use of a spin column for both purification and concentration.

11. Add 525 µL of 100% ethanol to the supernatant for a total volume of 825 µL. Place 700 µL into an RNeasy™ spin column and centrifuge >8000 r.p.m. for 1 minute.

12. Discard the flow-through and load the remainder of the sample on the column and centrifuge again @ >8000 r.p.m. for 1 minute.

13. Load 700 µL of buffer RPE and centrifuge again @ >8000 r.p.m. for 1 minute to wash the sample. Discard the flow through. Centrifuge again @ >8000 r.p.m. for 1 minute to dry the column.

14. Add 50 µL of RNase free water and centrifuge again @ >8000 r.p.m. for 1 minute to elute RNA from the membrane. At this stage, the sample can be analyzed using the Nanovue for concentration and a small subset (~2 µL can be loaded on to a gel for an analytic look at it.)

15. In order to precipitate the RNA, add 38.5 µL of nuclease free water, 1.5 µL of GlycoBlue and 10.0 µL of 3M sodium acetate pH 5.5 and 150 µL of isopropanol.

The GlycoBlue is a co-precipitant and it consists of glycogen that has been covalently linked to a blue dye (at this moment I do not know the name of the dye nor its chemical structure, I do know that it is pH responsive and will turn pink when a solution is alkaline). This serves two purposes, it adds mass to the nucleic acid in order to facilitate precipitation, it is capable of precipitating nucleic acids at concentration on the order of 1ng / mL. In addition to making the nucleic acid pellet have more mass to facilitate easier pelleting it also makes the pellet more visible. Thus far in subsequent reactions, the GlycoBlue does not interfere with any downstream applications, save one, it will interrupt the ABI sequencing performed by Genewiz.

16. Precipitate the RNA by chilling the mixture to -80 °C.

In order to get nucleic acids to aggregate for pelleting we need to neutralize charge by adding a salt (Sodium Acetate) and an alcohol (Isopropanol), we need

to lower to temperature to facilitate aggregation and we need time for all of this to happen. I have found that 30 minutes at -80 °C will freeze the samples, in order to facilitate more aggregation, I remove samples after 30 minutes thaw them and place them back into -80 °C for another 30 minutes.

17. Pellet the RNA by centrifugation for 30 mins @ 20 000g @ 4 °C in the refrigerated Eppendorf micro-centrifuge. Carefully pipette the liquid from the tube and let the pellet air dry for 10 minutes.
18. Re-suspend in 5 µL of 10mM Tris (pH 8.0)
19. Pre-run a 15% TBE-urea gel at 200 V for 15 mins in 1X TBE.
20. Add 5 µL of 2x denaturing loading buffer to each sample. Prepare the size selection markers (enough for two lanes) by adding 0.2 µL of 10 µM lower marker (18 nucleotide ssRNA marker) and 0.2 µL of 10 µM upper marker (42 nucleotide ssRNA marker) to 9.6 uL of 10 mM Tris (pH 8.0). Prepare ladder sample by adding 0.2 µL of 10bp ladder (1 µg/ µL) to 9.8 µL 10 mM Tris (pH 8.0)
21. Denature samples for 90 secs @ 80 °C
22. Flush wells of gel IMMEDIATELY prior to loading, if this is done too early on, urea may settle back into the wells and the samples may not run as efficiently. Load samples using gel loading tips, ensure that there is no overflow between lanes. It is important to use an even pressure when pipetting samples into the lane as uneven pressure may cause spill over into other wells as the force may be too vigorous and the sample is displaced from the well.
23. Run gel @ 200 V for 70 mins.

24. Stain with 1X SYBR Gold (2 µL into 20 mL of 1X TBE) for 3-5 minutes and visualize using dark illuminator (visible blue light LED illuminator).

If you intend to take a picture of the gel with a UV transilluminator ensure that the gel is exposed for the minimum amount of time possible. The UV light is very powerful in the picture taking apparatus and it will crosslink your sample. This will make downstream processing less efficient and it will reduce the integrity of your sample. The same principle applies for the UV transilluminator used for cutting gel slices. Although the wavelength of light used is higher, it is still strong enough to cross-link your sample or degrade if exposed to the light for prolonged periods. Using the dark illuminator is HIGHLY recommended, the dark illuminator uses visible blue light that excites the fluorophores without damaging the oligonucleotides.

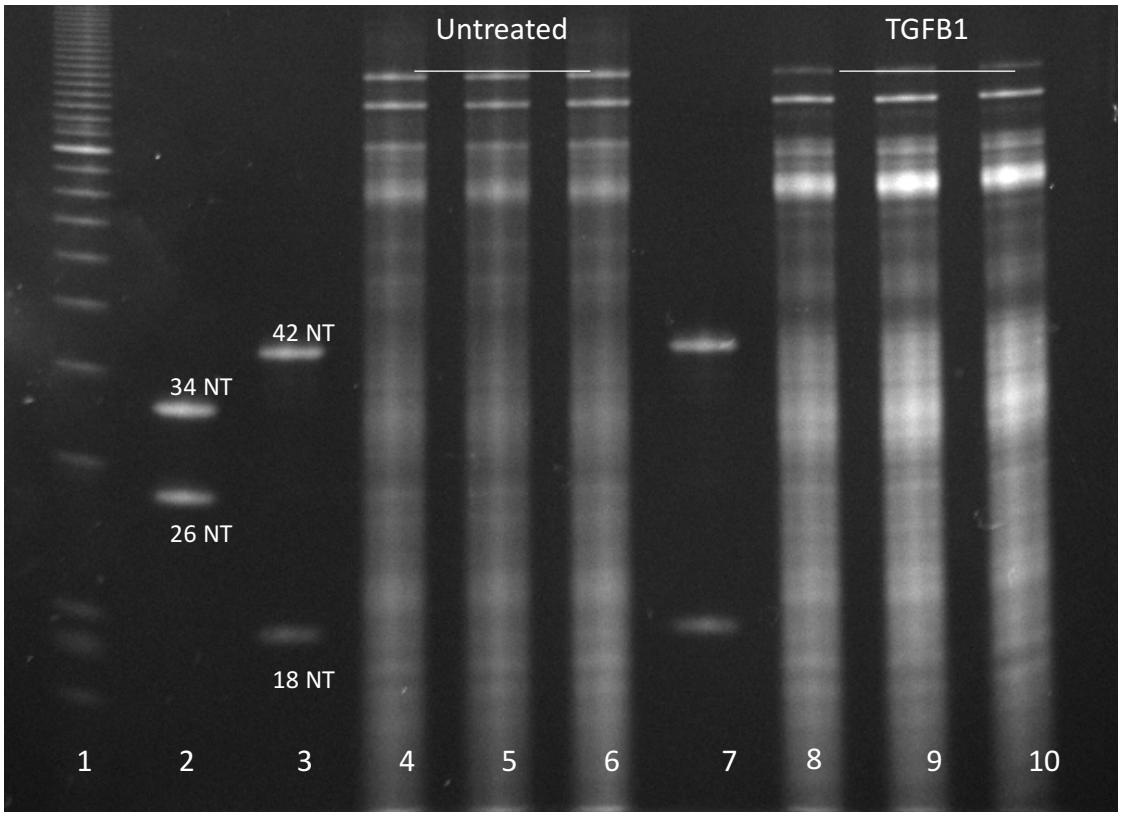

Figure 9. Picture of RNA gel showing the banding pattern of Rnase I digested ribosomes. Lane 1: Ladder. Lane 3 and 7: size selection markers, upper band (42 Nucleotides (NT)) lower band (18 Nucleotides (NT)). Lane 2: Previous Size selection markers 34 and 26 nucleotides (NT). Lane 4, 5, 6: untreated ribosome protected fragments. Lanes 8, 9, 10 TGFB1 treated ribosome protected fragments.

1.2.6 Gel Extraction:

25. Using a fresh razor or scalpel blade cut between the two size markers, the RNA should be contained between the two size markers. Use a fresh razor blade for each gel slice in order to avoid cross contamination of the samples. Place each gel slice into a new, clean micro-centrifuge tube.

ALTERATION:

As recommended by the protocol the two options for gel extraction are either over-night extraction using an extraction buffer or a rapid gel extraction whereby the polyacrylamide is physically disrupted and the nucleic acid extracted from

it. I have not extensively tested the rapid extraction but as it relies on more reagents and more consumables it was not considered a viable option. Instead of either of these techniques, a passive dialysis method has been implemented. This method relies on the principle that the force generated by the concentration gradient should be kept at a maximal level if the extraction buffer is exchanged frequently. The nucleic acid is at a high concentration in the gel but a low concentration in the buffer, as the nucleic acid equilibrates the rate at which it leaves the gel eventually slows, the idea is to remove the buffer before it approaches saturation levels of nucleic acid thus maintaining a maximal transfer rate from the gel into the buffer.

26. Add 100 µL of RNA extraction buffer to the gel slice, ensure that the gel slice is completely covered. Allow elution for 30 mins, then remove the 100 µL of buffer, place into a new non-stick micro-centrifuge tube. Repeat 3 more times for a total volume of 400 µL.

It is also a good idea to cut the gel slice as thinly as possible and into quarters, the more surface area available for nucleic acid exchange, the better the results will be.

27. Precipitate RNA by adding 1.5 µL of Glyco Blue, mixing well and then adding 500 µL of isopropanol. Recover RNA as described in steps 16 and 17.

1.2.7 De-phosphorylation reaction:

This step is the most salt sensitive step of the protocol, the enzyme T4 Polynucleotide Kinase is sensitive to high salt concentrations, it reduces the efficiency with which it operates. If for some reason this reaction is not

performing efficiently it may be wise to do an ethanol wash of your pellet before proceeding with this step.

28. Re-suspend the size selected RNA in 10 µL of 10 mM Tris (pH 8.0) and transfer to a new clean non-stick micro-centrifuge tube.

29. Prepare the de-phosphorylation reaction by adding 33 µL of nuclease free water to each sample. Denature samples for 90 secs @ 80 °C the equilibrate to 37 °C.

30. Set up each reaction as follows:

Table 1. Reaction set up for 1 reaction of de-phosphorylation of RNA.

Component	Volume (µL)	Final Conc.
RNA sample	43.0	
T4 PNK Buffer	5.0	1X
SUPERase.In	1.0	20 U
T4 PNK	1.0	10 U

31. Incubate the samples @ 37 °C for 1 hour then denature enzyme @ 70 °C for 10 minutes.

Precipitate the RNA by adding 39 µL of nuclease free water, 1.0 µL of GlycoBlue and 10.0 µL of 3 M sodium acetate, mixing well, then add 150 µL of isopropanol. Recover RNA as described in steps 16 and 17.

1.2.8 Linker Ligation:

(Timing: 4 hours)

32. Re-suspend the dephosphorylated RNA in 8.5 µL of 10 mM Tris (pH 8.0) and transfer to a clean non-stick micro-centrifuge tube.

33. Add 1.5 µL of pre-adenylated linker (0.5 µg μL^{-1}), denature @ 80 °C for 90 secs.

Set up the following reaction per sample:

Table 2. Reaction set up for 1 reaction of linker ligation of RNA.

Component	Volume (µL)	Final
RNA and linker	10.0	
T4 Rnl2 buffer (10X)	2.0	1X
PEG 8000 (50% wt/vol)	6.0	15% (wt/vol)
SUPERase.In (20U μL^{-1})	1.0	20 U
T4 Rnl2 (10 U μL^{-1})	1.0	10 U

34. Incubate the reaction for 2.5 hours @ room temperature.

This is one of the most pivotal parts of the protocol as in this step we ligate a known sequence of nucleotides, form henceforth it will be referred to as the **linker**, onto the unknown fragments. The linker forms the scaffold from which we build the corresponding parts of the library. From this addition, all subsequent parts of the library are built using various techniques as we seek to mimic the adapters used in Illumina™ sequencing. If this reaction does not work or is at a very low efficiency the subsequent reactions are impossible or they will not work to an appreciable degree. This reaction uses two key characteristics to accomplish a clean ligation of the unknown RNA to the known marker sequence, the removal of the 3' phosphate group by enzyme T4 polynucleotide kinase performed by the previous step and the ability of the enzyme T4 RNA ligase 2 truncated (T4 Rnl2) to join a ssDNA with a 5' adenylyl (5'adenine with a pyrophosphate group) group to the 3' OH end of ssRNA.

The ligation reaction takes place in the absence of ATP pushing the

reaction in the direction of joining the 5' end of the ssDNA strand to the 3' end of the ssRNA. One other point to note, although the pre-adenylated marker is expensive do not be tempted to use it in a lower concentration, it **MUST** be in excess of your RNA fragments as this increases the odds significantly that the enzyme will join a linker onto an RNA fragment, the enzyme cannot join two linker molecules as the 3' end of the linker is blocked by an NH_2 group.

It is feasible that the enzyme can join two ssRNA strands if one has a 5' phosphate group attached, however these reactions should be rare in occurrence given the excess of the linker and the concatenated ssRNA molecules will migrate to a different part of the gel. Last but not least, the use of PEG 8000 aids in the concentration of the reaction.

Enzymatic reactions require time and space, molecules have to be in the same physical space for the reaction to occur and there needs to be time for the reaction to occur. PEG 8000 is hydrophobic and when placed into the reaction mixture it pushes other molecules physically closer together it increases the concentration of the reaction which facilitates a more efficient reaction.

ALTERATION:

If the starting amount of RNA was below 1 µg then add 2 µL of T4 Rnl2 enzyme to the reaction mixture.

35. Add 338 µL of water, 40 µL of 3 M sodium acetate (pH 5.5) and 1.5 µL GlycoBlue to each reaction, mix well. Then add 500 µL of isopropanol. Recover RNA as described in steps 16 and 17.
36. Separate the ligation reactions by polyacrylamide gel electrophoresis as described in steps 19 – 23.

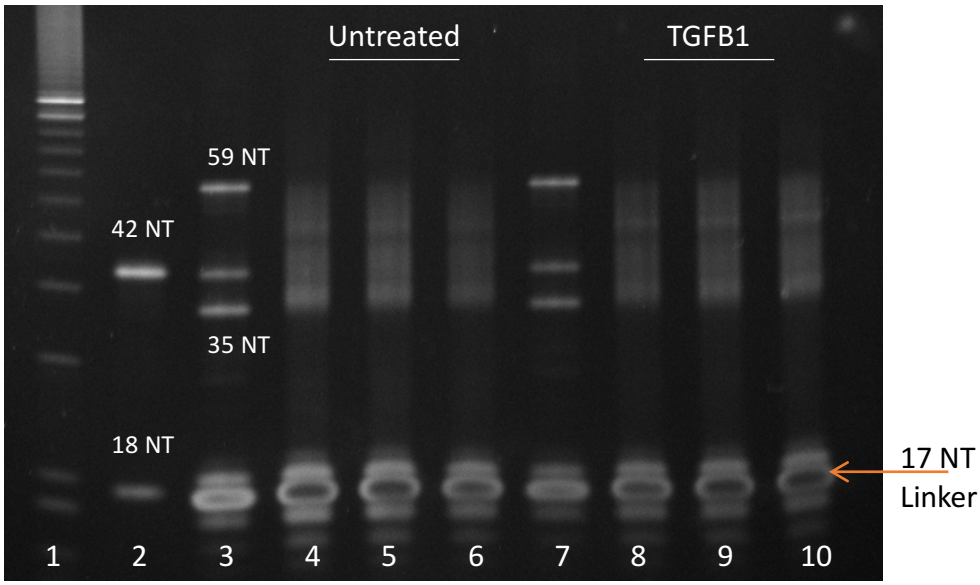

Figure 10. Picture of a gel showing successful ligation products as well as positive controls using the size marker. Lane 1: Ladder. Lane 3 and 7: Ligated size selection markers, upper band (59 Nucleotides (NT)) lower band (35 Nucleotides (NT)). Lane 2: Size selection markers 18 and 42 nucleotides (NT). Lane 4, 5, 6: Untreated, ligated ribosome protected fragments. Lanes 8, 9, 10 TGFB1 Ligated treated ribosome protected fragments.

37. Excise the ligation product band and place each gel slice in a clean non-stick micro-centrifuge tube. Recover RNA as described in step 26 and 27.

1.2.9 Reverse Transcription:

(Timing: 2 hours)

This reaction was previously the rate limiting step in the protocol as I have thought it was inefficient and slow to give a sufficient quantity of product. However, I have discovered that this reaction, although inefficient, produces product even with low (< 2 pmol) starting material quantities. The product band that is produced is often difficult to visualize but using the amount of product produced in subsequent reactions produces the expected results. In order to

increase the efficiency of the reaction I have altered it by adding 1.25 more pmoles of primer and 200 more units of enzyme.

ALTERATION:

Instead of 2.0 µL of primer use 3 µL and instead of 1.0 µL of enzyme use 2.0.

38. Re-suspend the ligation product in 10 µL of 10 mM Tri (pH 8) and transfer to a clean PCR tube.

39. Add 3.0 µL of reverse transcription primer that was reconstituted to 1.25 µM. Denature for 2 mins @ 80 °C. Place on ice until reverse reaction is set up.

Set up the following per sample:

Table 3. Reaction set up for 1 reaction of reverse transcription of RNA into ssDNA.

Component	Volume (µL)	Final Concentration
Ligation and primer	13.0	
First strand buffer (5X)	4.0	1X
dNTPs	1.0	0.45 mM
DTT	1.0	4.54 mM
SUPERase.In	1.0	20 U
SuperScriptIII(200UµL^{-1})	2.0	400 U

40. Run reverse reaction at 48 °C for 30 mins.

41. Hydrolyze the RNA by adding 2.2 µL of 1N NaOH to each reaction, incubate @ 98 °C for 20 mins. The GlycoBlue will turn pink.

Hydrolysis of ssRNA using high temperatures and alkaline conditions occur because the 2' OH^- group is susceptible to nucleophilic attack under alkaline conditions and high temperatures but because DNA has a 2' H^+ it is not susceptible to alkaline hydrolysis and thusly will not be degraded as the RNA will.

42. Add 20 µL of 3 M sodium acetate (pH 5.5), 2.0 µL GlycoBlue and 156 µL of nuclease free water to each reaction, mix well then add 300 µL of isopropanol. Recover ssDNA as described in steps 16 and 17.

43. Separate the reverse-transcription products from the unextended primer by polyacrylamide gel electrophoresis as described in steps 19 – 23.

The reverse transcription primer has two hexaethylene glycol spacers referred to in the commercial oligonucleotide business as (SpC-18). The spacer consists of an 18 atom long chain of 12 carbon and 6 oxygen molecules that adds flexibility to the RT primer[40]. In addition to adding flexibility for the subsequent circularization step, the hexaethylene glycol spacer also acts as a stop point in the PCR step as it contains no nucleotides. The lack of nucleotides ensures that the polymerase cannot polymerize through the relatively large stretch of non-nucleotide atoms. It is my opinion that the hexaethylene spacers also causes an inordinate amount of smearing in 15% polyacrylamide-urea gels. When run on a gel to separate the reverse-transcription reaction there will be background smearing in the lanes, this occurs in the lane that contains only RT primer as a positive control. I believe it is a property of the primer chemistry and not necessarily poor gel mixing/polymerization, nucleotide degradation or any other part of the experiment.

44. Separate the reverse transcription products from the unextended primer by polyacrylamide gel electrophoresis as described in steps 19 – 23. As a control, prepare one sample with 3.0 µL of RT primer (1.25 µM), 2.0 µL of 10 mM Tris (pH 8) and 5.0 µL of 2X denaturing sample buffer.

45. Extract the DNA from the gel slices by using the methods described in step 26. Note it is necessary to use the DNA extraction buffer and NOT the RNA extraction buffer.

46. Precipitate the DNA by adding 1.5 µL of GlycoBlue, mixing well and then adding 500 µL of isopropanol. Recover DNA as described by steps 16 and 17.

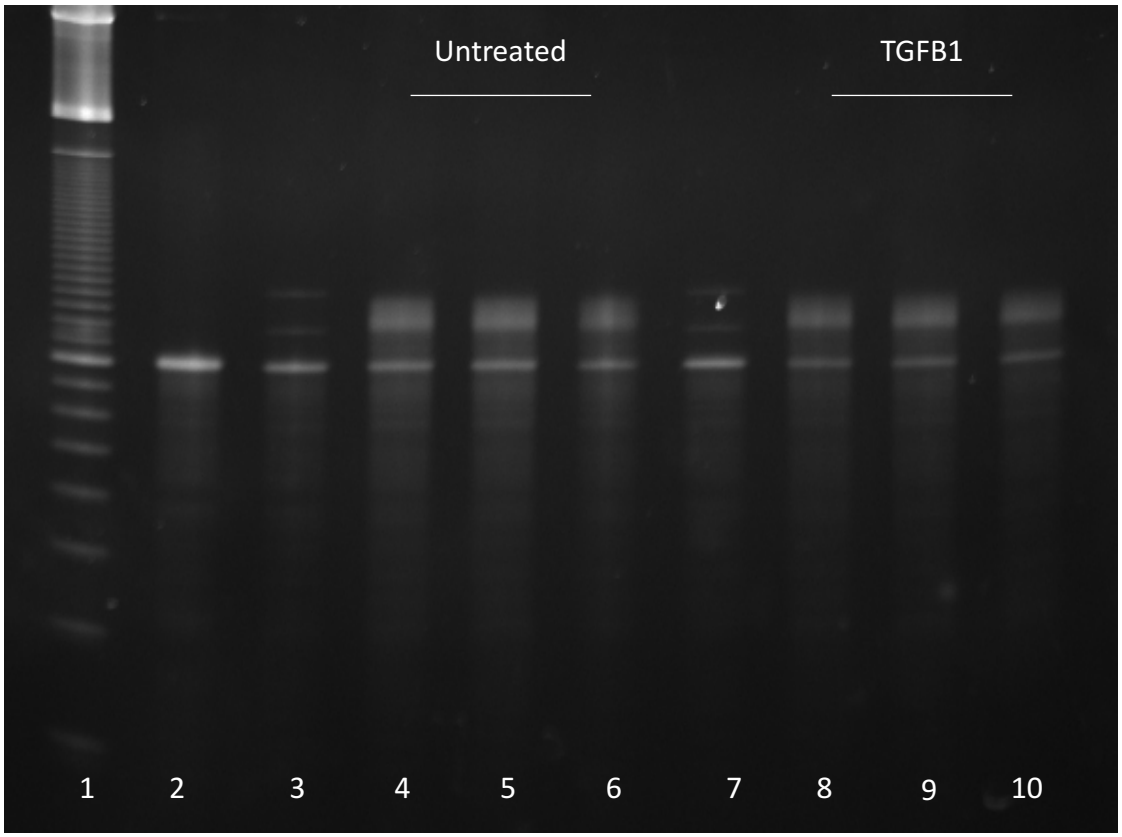

Figure 11. Picture of gel showing successful reverse transcription reactions. The size markers were used as a positive controls. Lane 1: Ladder. Lane 3 and 7: Reverse Transcribed size selection markers, upper band (159 Nucleotides (NT)) lower band (135 Nucleotides (NT)). Lane 2: RT Primer ~100 nucleotides (NT). Lane 4, 5, 6: Untreated, reverse transcribed ribosome protected fragments. Lanes 8, 9, 10 TGFB1 treated reverse transcribed ribosome protected fragments.

1.2.10 Circularization:

(Timing: 1.5 hours)

45. Resuspend the reverse transcription products in 15.0 µL of 10 mM Tris (pH 8) and transfer to a PCR tube.

Prepare the circularization reaction for each sample as follows:

Table 4. Reaction set up for 1 reaction of circularization of linear ssDNA into circular ssDNA.

Component	Volume (µL)	Final Conc.
Template	15.0	
CircLigase buffer (10X)	2.0	1X
ATP (1mM)	1.0	0.05 mM
$MnCl_2$ (50 mM)	1.0	2.50 mM
CircLigase	1.0	100 U

The CircLigase enzyme is capable of working with low amounts of starting template (1 pmol), I would strongly recommend using the control template provided with the enzyme as a positive control and to test whether or not the enzyme is active. With regards to the positive control, it is provided as ssDNA and if it circularizes the is an upward shift in how it migrates in the gel. Circularization is a **KEY** part of the procedure if it does not work the library will not be built correctly. As a positive control the correct fragment sizes will only be seen if circularization has successfully occurred.

1.2.11 PCR amplification and Index Addition:

(Timing: 3 hours)

In order to generate enough material to put onto the Illumina machine one has to do PCR amplification and in order to run multiple samples on the same lane of a flow cell (referred to as multiplexing) one has to add indexing barcodes to the samples in order to be able to distinguish the samples from each other. In order to make a mixture of the correct concentration to load onto the Illumina

flow cell the concentration of your final product has to be greater than 2 nM. The

Table 5. Component volumes of each reaction component need to set up one PCR reaction for indexing the libraries.

final concentration of the working mixture is made to 2 nM thus your sample has to be greater than 2 nM in order to be at the right concentration in the final mixture. For optimal cluster generation, it is recommended that the concentration loaded on the flow cell is between 8 pM – 10 pM.

46. Set up 5 PCR strips and transfer a 16.7 µL of the PCR master mix

Component	Volume (µL)	Final Conc.
Phusion HF buffer (5X)	20	1X
dNTPs (10 mM)	2.0	0.2 mM
FWD primer (10 µM)	5.0	0.5 µM
Reverse Index primer (10 µM)	5.0	0.5 µM
Circularized DNA template	5.0	
Nuclease free Water	62	
Phusion polymerase	1.0	2 U

into one tube in each strip.

47. Perform the PCR with varying number of cycles, do this by starting the program and removing PCR tubes as the cycle number is reached. Be sure to let the extension be completed before removing the tube. Remove tubes after cycle 6, 8, 10 and 12 leaving the last tube in for the completion of cycle 14.

Table 6. Temperature and times for the PCR program set up in order to run the PCR reactions.

Cycle Number	Denature	Anneal	Extend
1	98 °C, 30 s		
2-15	98 °C, 15 s	65 °C, 15 s	72 °C, 15 s

48. Add 3.3 µL of 6X non-denaturing loading buffer to each reaction. Prepare a ladder sample with 0.2 µL of 10 bp ladder, 9.8 µL of 10 mM Tris (pH 8) and 2.0 µL of 6X non-denaturing loading buffer.

49. Separate the PCR product using an 8% non-denaturing gel, load amplification reaction from the same sample into adjacent well for direct comparison.

50. Run the gel @ 180 V for 40 mins. Stain the gel as described in step 24.

51. Excise PCR product bands that are prominent but with little accumulation of the re-annealed partial duplex library products. Place these gel slices into a new, clean non-stick micro-centrifuge tube.

52. Recover DNA from gel slices as described in steps 46 – 47

53. Re-suspend the library in 15.0 µL of 10 mM Tris (pH 8). It may be stored indefinitely at 4 °C or -20 °C.

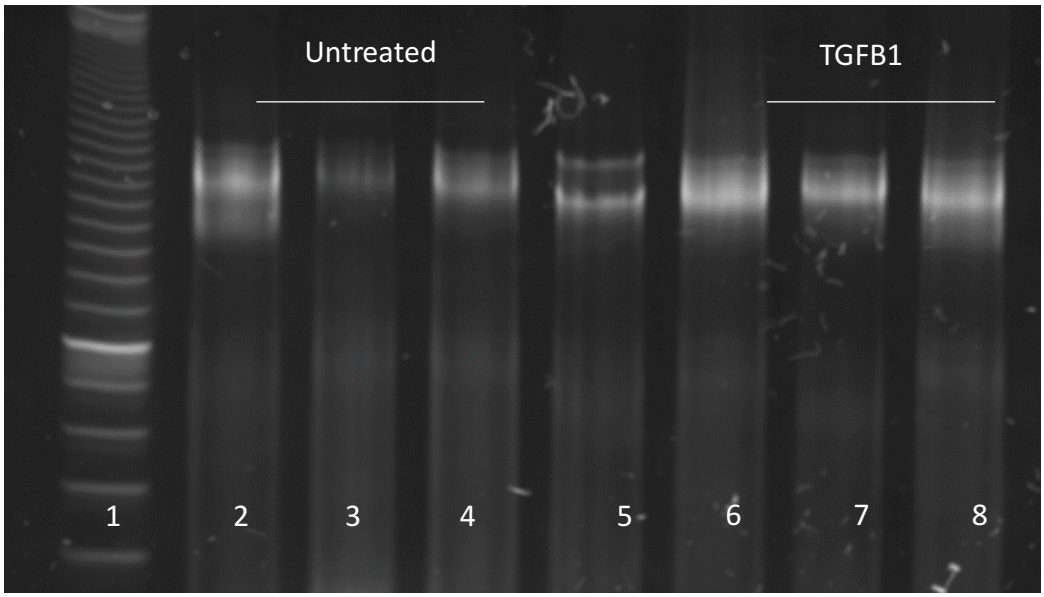

Figure 12. Lane 1: Ladder. Lanes 2 ,3 and 4: Indexed library for Untreated Sample #1, 2, 3. Lane 5 Indexed size marker positive control. Indexed library for TGFB1 samples. Lanes 6,7, and 8 are indexed libraries.

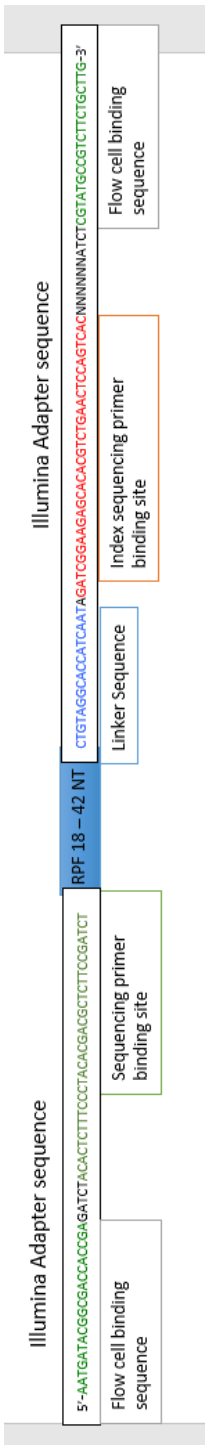

Figure 13. Schematic diagram of the finished Illumina sequencing compatible indexed library. All functionally important parts of the library are highlighted.

1.3 CONCLUSION

Following the outlined steps above for the implementation of the ribosome profiling protocol yielded highly reproducible results and was the tool needed to get a global snap shot of translation changes that occur when cells go through EMT. Libraries generated are stranded, *i.e.* you can determine if the mRNA read came from the plus (+) or minus (-) strand of DNA and can be sequenced with Illumina single read 51 (SR51) sequencing runs. Paired end sequencing will NOT work as the primers are not designed for that.

If libraries are run on a HiSeq 4000 machine, then one can expect around 350 – 380 million reads in total for one lane.

CHAPTER 2

RIBOSOME PROFILING REVEALS DISTINCT BUT COORDINATED TRANSLATION PROGRAMS IN EMT

2.1 OVERVIEW

The epithelial to mesenchymal transition occurs during normal development in vertebrate embryos[25,41]. It was first described by Dr. Elizabeth D. Hay using the chick embryo development system and been shown to be a naturally occurring process[12–14]. It has also been proposed that the re-activation of the EMT process during adulthood leads to disease states particularly[16,42]. These ideas were discussed in more detail during the general overview.

Building off of the previous work done in implementing the ribosome profiling methodology it was now possible to apply it to cells that are going through or those that have undergone EMT using the NMuMG system[43]. The NMuMG EMT system is a TGFβ driven system where cells undergo EMT in response to the addition of TGFβ and perhaps more intriguingly the EMT can be reversed by the removal of the TGFβ stimulus[23]. The reversal of the EMT phenotype by the removal of the stimulus allows for the interrogation of how the cells revert to an epithelial state.

Despite being only one of the main signaling molecules to initiate the EMT program the TGFβ driven NMuMG system is more amenable to the Ribosome Profiling technique as it offers a ready supply of material to work with, a controlled system in which one can efficiently trigger EMT as well as apply small molecule inhibitors of various key components of the TGFβ signaling pathway. Given the previously mentioned commonalities in the other EMT systems (Wnt driven or Hypoxia driven) it is possible to then test observations made in the NMuMG system in the other EMT systems.

Lastly there also exists the possibility to test whether or not any of the molecular players (biomarkers) that are discovered using the ribosome profiling technique can also be observed in human tissues. Using samples from patients of both cancerous and normal tissues we can get a look into how amenable the observations made in our model system to observations made of human samples. These observations would lend credence to the data and the possibility that upon further investigation that these biomarkers may be used for some clinical applications.

Data generation acknowledgement: In the following figures multiple people contributed data for figure generation. ***Drs. Brittany Carson and Jennifer Feenstra*** contributed all of the immunofluorescence images and western blots for main text and supplementary figures fourteen (14) and fifteen (15). **All** statistical and bioinformatics analyses were performed by ***Dr. Matthew Parks***. All Immunohistochemistry (IHC) data of human samples were generated by ***Varsha Prakash***. ***Chad Kurylo*** performed polysome profile analyses.

2.1.1 Introduction

Ribosome catalyzed protein synthesis is central to cellular homeostasis, growth and proliferation and is widely recognized to be a driving aspect of uncontrolled cancerous cell growth and proliferation[44,45,46,47]. It has been recently shown that the non-proliferative, Epithelial-to-Mesenchymal transition (EMT) is considered central to metastatic disease[17,48,49] requires unregulated ribosome biogenesis (Prakash *et al*.). These findings suggest that aspects of protein synthesis mechanism that have yet to be explored may be essential to reprogramming the cell for the mesenchymal state.

Though there has been recent progress in understanding metastatic spread[4,5,50], the exact mechanisms underlying the dissemination of non-proliferating cells capable of colonization secondary organ systems remain unclear. Cells that have undergone EMT lose junctional adhesions and gain of mesenchymal cytoskeletal proteins to acquire pro-invasive properties that enable them to leave the primary tumor site[16,26,51]. During EMT the cell cycle is arrested with concomitant activation of pro-mesenchymal gene expression driven by transcription factors including, but not limited to, Snail, Smads, Twist and β-catenin[18,19,52], which are strictly regulated at both a transcriptional and a translational level[30,53].

By contrast, the global translational landscape underlying gene expression during execution of the EMT program is less well understood. Here, we modify the ribosome profiling method[33] to obtain genome-wide snapshots of changes in mRNA, tRNA and rRNA utilization pre- and post-TGFβ-mediated EMT. We leverage these investigations to identify and validate new biomarkers that have not previously been associated with EMT, including key components of the translation apparatus. The observed upregulation of the pro-mesenchymal translational program occurs during a period of reduced mTOR signaling. These changes were exacerbated by further reduction of mTOR signaling or by preventing new rRNA from being synthesized. Confirming the pervasive nature of the translational reprogramming during EMT, we further demonstrate that cells committed to the mesenchymal state are specifically sensitive to translation elongation inhibitors. These findings inform on potential therapeutic strategies for targeting and preventing the dissemination of metastatic disease.

2.1.2 Ribosome Profiling Reveals Novel Markers of EMT

To investigate the translational landscape underpinning EMT, we employed the NMuMG mouse mammary epithelial cell line, an extensively interrogated TGFβ-inducible EMT model system. As expected (Prakash *et al*), this system efficiently adopted a pro-migratory, mesenchymal phenotype after 48 hours of TGFβ treatment, accompanied by a marked induction of rDNA transcription and a global reduction in protein synthesis. To examine changes in products of all three RNA polymerases (mRNA, tRNA and rRNA) during EMT, we performed a modified ribosome profiling method, in which no nucleic acid bio-type was subtractively removed (**Fig.S14b**).

Our analysis of gene expression regulation during EMT began with comparisons of change in total mRNA transcription within NMuMG cells pre- and post-EMT and ribosome protected fragments generated from the actively translated mRNA pool (**Fig.14a**). As previously shown[17,28], NMuMG cells treated with TGFβ exhibited a loss of the epithelial markers, including E-cadherin (Cdh1) and coxsackie and adenovirus receptor (Cxadr), simultaneous with a gain in expression of the mesenchymal markers N-cadherin (Cdh2) and Snail (Snai1) at both the mRNA and protein level (**Fig.14a; Fig.S14c**). In total, we detected 3142 genes to be differentially expressed on a transcriptional, translational or transcription and translational level after 48 hours of TGFβ treatment (FDR < 0.05), which represents approximately 25% of all genes identified. Beyond the multitude of established EMT markers, we observed changes in a number of genes that had not been previously associated with EMT, including Long Non-coding RNAs (lncRNAs), small nuclear and nucleolar RNAs (snRNAs and snoRNAs, respectively), microRNAs (miRNAs) processed pseudogenes as well as mRNA transcripts (**Fig.S14b**).

To validate the association of these genes as pro-mesenchymal/EMT markers and the sensitivity of the ribosome profiling approach, we first probed the modestly increased transcriptional and translational expression of mRNA transcripts not previously associated with the EMT program. Cttnbp2nl was chosen as it is associated with stress fibers and cytoskeletal structures that accompany other pro-migratory and invasive cytoskeletal rearrangements[54] (**Fig.S14e**). Eif6 was chosen because they are directly involved in ribosome biogenesis[55]. Eif6 (also known as Tif6) binds the large ribosomal subunit to assist ribosome maturation and to prevent premature subunit association[56,57]. The observed transcriptional and translation changes of all markers were confirmed by immunofluorescence and immuno-blot analysis (**Fig.14b**). All markers were also observed to respond similarly in a second TGFβ-inducible EMT model, the Py2T system (**Fig.14c**)[58,59], and after hypoxia-induced EMT in estrogen receptor-positive human MCF7 cells (**Fig.14d**) as evidenced by the loss of CDH1 and a gain of SNAI1 expression and nuclear localization (**Fig.S14d**). All markers also exhibited increased expression in the triple negative MDA-MB-231 breast cancer cell line compared to the estrogen receptor positive (ER+) MCF7 (**Fig.14e**)[60]. Eif6 was further shown to be highly expressed in the migratory neural crest cell population which have undergone Wnt-driven EMT[25,61] that also exhibits high rRNA expression(rDNA paper) (**Fig.14f**). These new EMT markers were further validated by examining normal and invasive human tumor tissue samples, where in each case, invasive tumor samples showed increased expression (**Fig.14g**). These findings demonstrate the power of the ribosome profiling method to identify a diversity of genes specifically associated with the EMT program that may serve as markers of tumor progression and metastasis.

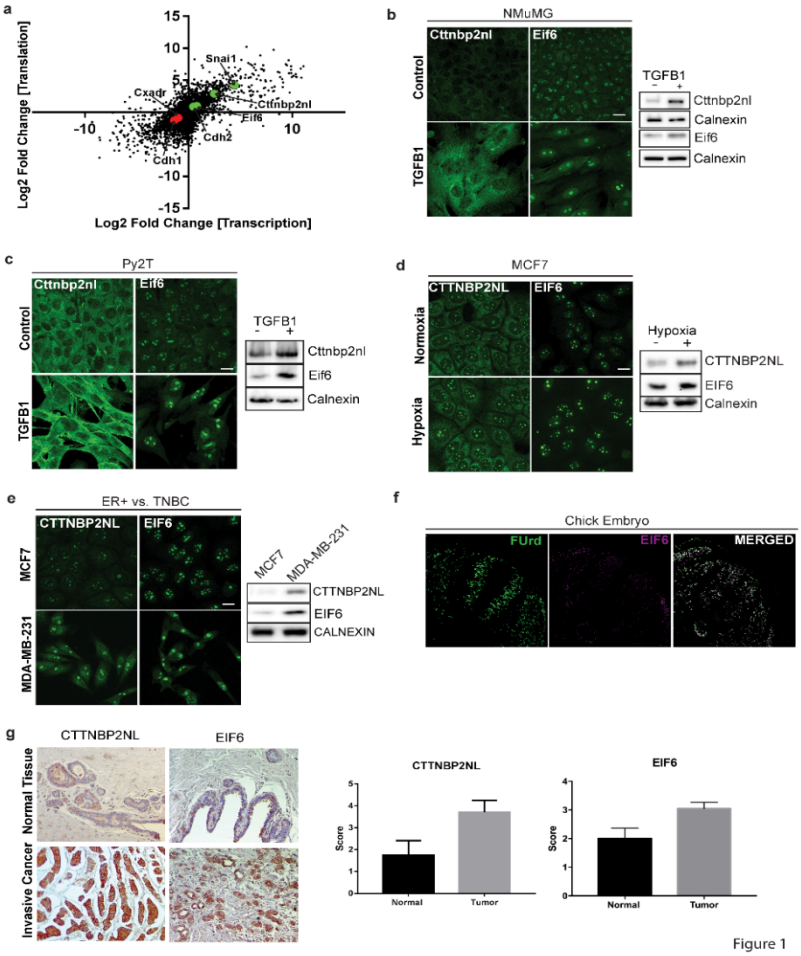

Figure 14. Ribosome profiling demonstrates translational regulation in EMT coupled with transcription and identification of novel EMT markers of translational regulation in EMT coupled with transcription and identification of novel EMT markers. (a) Scatter plot of genes that are differentially expressed in response to TGFβ and their corresponding changes on a transcription level (x axis) and on a translation level (y axis) FDR <0.05 **(b)** Changes in protein level expression and localization of transcriptionally and translationally co-regulated genes, Cttnbp2nl (green), and Eif6 (green) in TGFβ induced EMT in NMuMG cells **(c)** Changes in protein level expression and localization of transcriptionally and translationally co-regulated genes Cttnbp2nl (green) and Eif6 (green) in TGFβ induced EMT in Py2T cells **(d)** Changes in protein expression of CTTNBP2NL and EIF6 in hypoxia (48 hours) inducible EMT in MCF7 cells **(e)** Comparison of expression levels of CTTNBP2NL (green) and EIF6 (green), in estrogen receptor positive cells (MCF7) and triple negative breast cancer cells (MDA-MB-231) **(f)** Chick neural crest delamination showing co-localization of eIF6 with FUrd in the migratory population of cells **(g)** Increased expression of newly identified markers, CTNNBP2NL and EIF6 in invasive breast cancer. Graphs depict scoring of stain intensity by two independent researchers. Immunofluorescence and western blots for panels (**b,c,d)** and e done by Dr. Brittany Carson. Immunofluorescence for panel (**f)** was done by Dr. Jennifer Feenstra and IHC for panel **(g)** was done by Varsha Prakash. Panel **(a)** was made using analysis from Dr. Matt Parks.

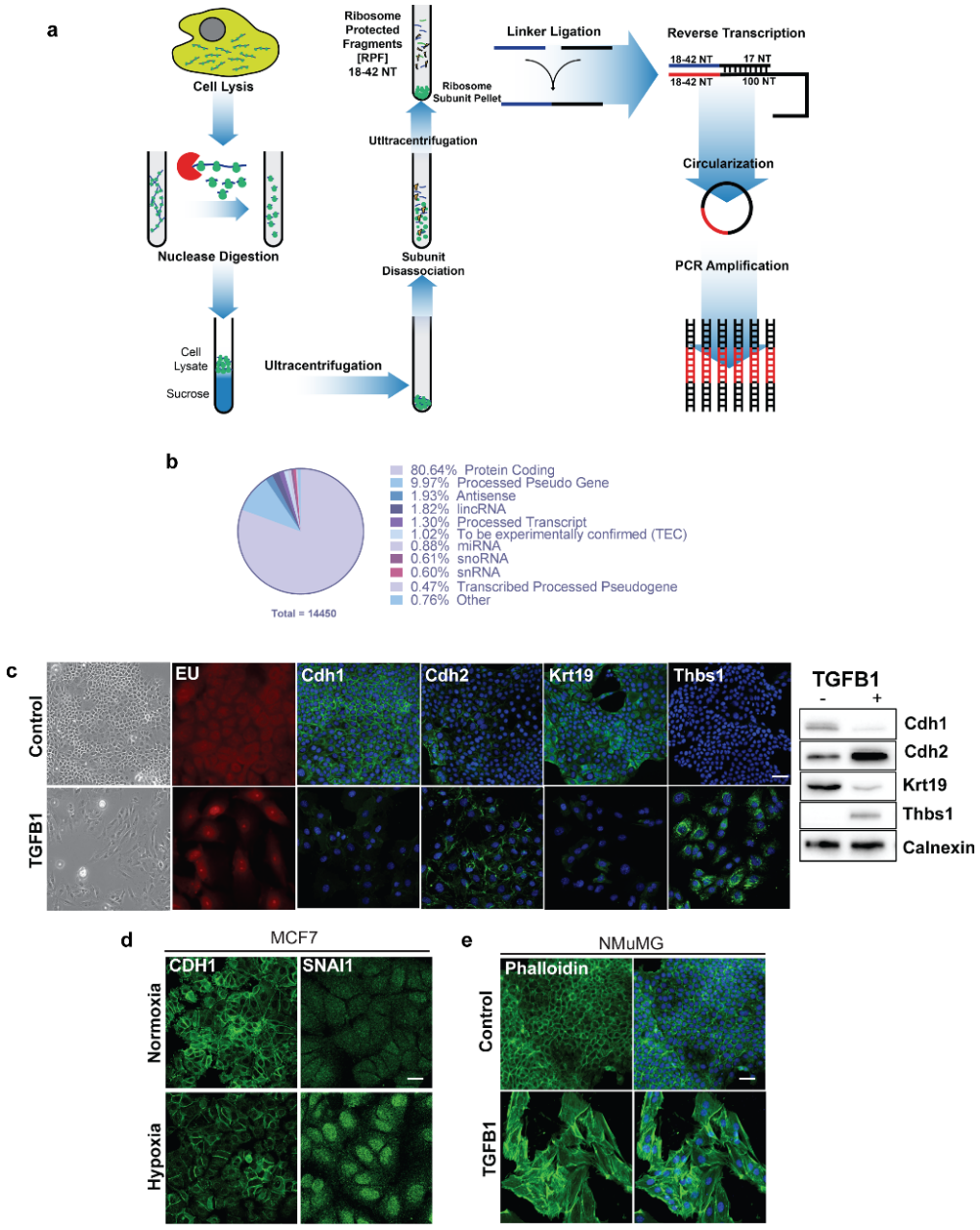

Figure S14. Ribosome Profiling Reveals Novel Markers of EMT (**a**) Schematic overview of modifications made to the existing ribosome profiling protocol (**b**) Graphical breakdown of all transcripts detected into respective categories (**c**) Phase contrast images demonstrating the morphological change that accompanies induction of EMT in NMuMG cells as well as changes in protein expression of established markers of EMT using both immunofluorescence and immunoblot (**d**) Loss of the epithelial marker E-cadherin and increase in nuclear accumulation of the transcription factor SNAIL in response to hypoxic conditions for 48 hours in MCF7 cells (**e**) Formation of stress fibers in response to induction of EMT and visualized by Phalloidin staining (green). Western blots and immunofluorescence of Cdh1, Cdh2 Krt19 and Thbs1for panel (**c**) as well as all of panel (**d**) and (**e**) were done by Dr. Brittany Carson. Panel **(b)** was made using analysis from Dr. Matt Parks.

2.1.3 Direct Observations of Pervasive Alterations in the Translation Machinery

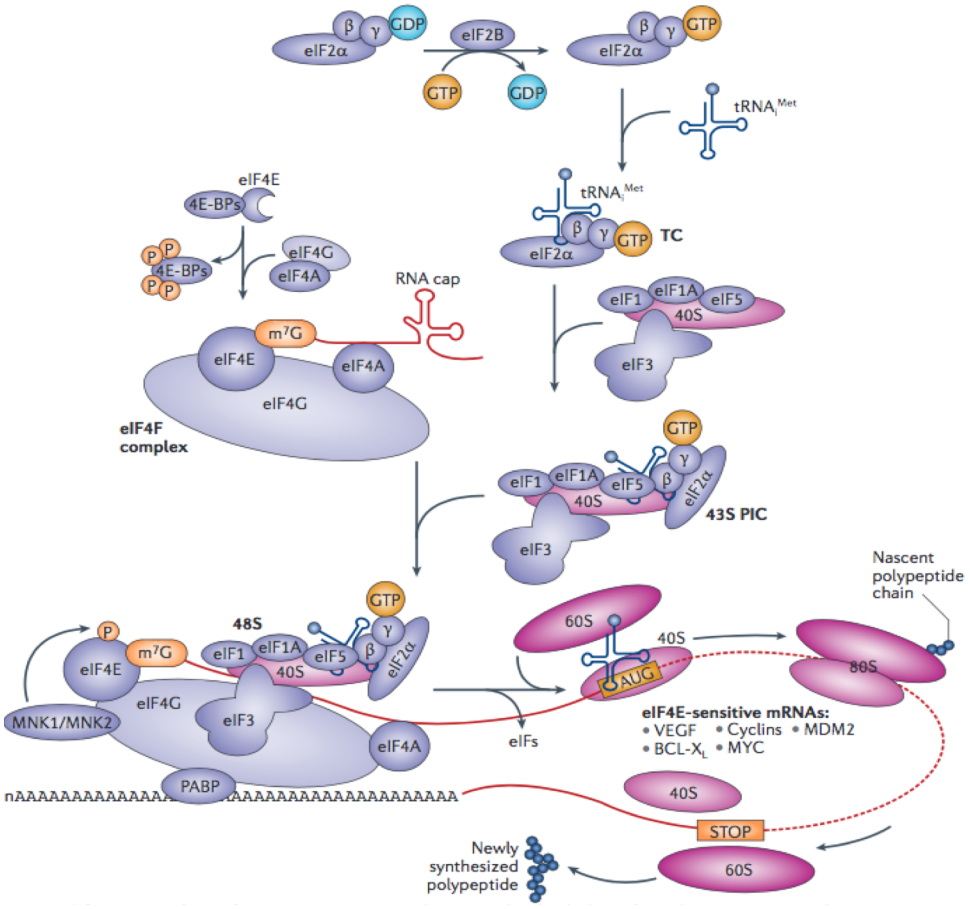

Figure 15. Schematic diagram showing the key players in the eukaryotic translation initiation complex. Canonical translation initiation is one of the most studied aspect of translation as it is thought to be the rate limiting step from translation. Diagram taken from Bhat et al. 2015 Nature Reviews Drug Discovery[65].

2.1.3.1 Canonical Translation Initiation versus Non-Canonical Initiation

Canonical or cap-dependent translation initiation utilizes the ^{7}methyl guanosine cap found on the 5' end of many mRNAs and is considered to be the rate-limiting step in protein synthesis[62] (**Fig. 15**). Translation initiation is a well-orchestrated process involving a multiplicity of protein factors **(Fig. 15)** as well as the hydrolysis of GTP molecules in addition to several phosphorylation events. Given the number of components and energy expenditure involved in cap-dependent translation there has been a significant amount of effort spend on deciphering the temporal sequence of events as well as the players involved in canonical translation initiation[53,63–66].

Canonical translation has been studied and implicated as being misregulated in the context of cancer[65,67–71]. We therefore examined a subset of translation initiation factors identified by ribosomal profiling to be translationally controlled post EMT including Eif3i, Eif3f and the ATP-dependent DEAD box RNA helicase Eif4a1 and its respective isoform Eif4a2. Given the recent data[72,73] that Eif3 can be involved in the specific translation of proto-oncogenic genes like c-Jun looking how components of the complex changed in response to EMT seemed prudent. Eif3i expression localization changed whereas Eif3f expression level modestly increased in line previous findings that increased expression is associated with a global reduction in translation[69,74].

A remarkable reduction of only one of the isoforms of the RNA helicase Eif4a1 was observed post EMT while a modest induction of Eif4a2 could be detected. This specific change in expression level of these two isoforms has previously been shown to be directly related to cell growth where the synthesis and translation of Eif4a1 is more efficient in proliferating cells compared to Eif4a2 which is translated in growth-arrested cells[75].

There has however been more evidence within the literature that there may be more pervasive uses of non-canonical or cap-independent initiation than previously thought. The most studied alternative form of translation initiation is the use of internal ribosome entry sites (IRES) of which the cricket paralysis virus (CrPV) is a well-studied model[76–79]. This form of initiation bypasses the use of the ^{7}methyl guanosine cap as well as the majority of the initiation complexes. As technological advances have been made, ribosome profiling in particular, we have been able to observe other manifestations of non-canonical or cap-independent translation in particular the use of unconventional start sites, upstream open reading frames and the repurposing of some components of the canonical translation machinery[34,35,63,71,80–82].

The implications of these observations are that translation initiation is more complex and varied than originally thought. It adds another layer of complexity to the translation control repertoire that can add multiple levels of protein variation from the same transcriptional background.

2.1.3 Translation Control is Pervasive in EMT

We further observed that approximately 1100 genes, or 30% of those identified as changed during EMT, exhibited a predominant increase or decrease in translation efficiency (**Fig. 16a**). These findings suggest pervasive changes in the translation program during EMT. Gene ontology (GO) analysis revealed that the three largest mRNA subsets displaying differential translation were integral membrane proteins as well as components of the cytoskeleton and translation machineries. The latter subset included most ribosomal proteins as well as specific translation initiation and elongation factors (Eif3k, Eif2s2, Eif5a Eif4a1, Eif3f, Eif1ad, Eif4g2, Eef1a1, Eef1b2, Eef1d and Eef1g). The

affected initiation factors comprise distinct components required for the small ribosomal subunit (40S) to locate the proper start site of protein synthesis. The affected elongation factors ensure rapid and efficient protein synthesis. eEF1a1 forms a ternary complex (TC) with aminoacyl-tRNA (aa-tRNA) and GTP and is responsible for the rapid and efficient delivery of aa-tRNA to the ribosome during active protein synthesis[83,84]. eEF1b2, eEF1d and eEF1g constitute the guanosine nucleotide exchange factor components for eEF1A and are thus responsible for maintenance of the TC pool. We validated these findings by performing targeted immunostaining and immuno-blotting studies on the modestly altered proteins Eef1b2, the main guanine nucleotide exchange factor (GEF) of TC and Gem, a putative GTP binding protein that has no known link to the translation machinery. The expression levels of both genes were confirmed in each *in vitro* EMT model system examined (**Fig. 16b; Fig S16a**). Strikingly, these studies reveal Eef1b2 to be nuclear localized after TGFβ treatment, suggesting the principle GEF component for TC is sequestered away from the translating ribosome in the mesenchymal state. The expression levels of EEF1B2 and GEM were also found to be similarly altered when invasive human tumor tissue samples were compared to normal tissue (**Fig.16a**,**c**). Remarkably, the nuclear localization of EEF1B2 was recapitulated in the invasive human tumor cell population.

2.1.4 Translationally Downregulated mRNAs Exhibit Features Characteristic of mTOR Control

To understand the underlying mechanisms giving rise to translation regulation during EMT, we evaluated the physical characteristics of the mRNA transcripts exhibiting translational control including the five prime (5') and three prime (3') untranslated region (UTR) length and 5' and 3' UTR folding energy.

In so doing, we found that genes showing reduced translation had shorter and less structured 5' and 3' UTRs (**Fig.16d; S16g**). These characteristics are typical in genes upregulated by mTOR signaling ("mTOR sensitive")[68], the majority of which are regulated by canonical, cap-dependent translation mechanisms[85]. This finding is consistent with the global reduction in protein synthesis evidenced in cells and in polysome profiles (**Fig.S16e; S16f**). Consistent with these findings, we observed that mTOR signaling was reduced post EMT evidenced by reduced expression of total mTOR kinase (shared between mTORC1 and mTORC2 complexes), mTORC2 mediated Akt phosphorylation, pS473, and Akt mediated phosphorylation specific mTORC1 mark pS2448, were all reduced post EMT. Further evidence of reduced mTOR signaling was displayed by reduced Rps6 phosphorylation and increased eIF2α phosphorylation in both NMuMG and MCF7 cells (**Fig.S16c,d**). That reduced mTOR is a conserved feature of the EMT program was further supported by expression studies of increased eIF2α phosphorylation in the chick developing embryo. The downregulation of mTOR likely contributes to the non-proliferative status of the mesenchymal state[68,86,87].

The inverse relationship between TGFβ and mTOR signaling was confirmed by serum-starving NMuMG cells in the absence and presence of TGFβ treatment. These investigations demonstrated that serum starvation further reduced mTOR signaling and intensified the EMT program as evidenced by morphological changes, induced Vimentin and N-cadherin expression and increased invasive capacity (**Fig.S16h**). These observations are consistent with observations that mTOR inhibition can induced a paused pluripotent stem cell-like state[22,48,88,89].

2.1.5 Translationally Upregulated mRNAs Exhibit Features Characteristic of Oncogenes

By contrast, the genes translationally induced by TGFβ were typified by longer and more structured 5' and 3' UTRs (**Fig. S16g**). Such features have been shown to be enriched in oncogenes implicated in tumor progression that exhibit non-canonical, cap-independent translation mechanisms[63]. In the present dataset, three of the 23 genes previously validated as *bona fide* tumorigenesis promoting, IRES-containing cellular mRNAs (Egr2, VEGFA, Jun,)[63] were observed to be translationally upregulated.

In addition to Gem and Vimentin (**Fig. 16a**), the translationally upregulated, pro-mesenchymal transcripts induced by TGFβ included integral membrane proteins (Adgrg1, Jak2, Klrg2, Slc4A1, Slc8a1, Slc25A25, Ttyh1), chromatin modifiers (Bahcc1, Clmn, Fam60a and SkiL), transcription factors (Cebpb, Csl, Foxp1, Foxo6, Fosb, Fosl2, Irf1, Junb, Klf10, Maf, Runx1t, Stat2/3), cytoskeletal proteins (Krt78, Synpo2l. Tubd1, Wipf2), RNA binding proteins and helicases (Rmrp, Stau1, Mov10, and Ddx3x/24/52) as well as specific translation factors, ribosomal proteins and regulatory factors of translation (eIF2AK4, eIF3F, eIF1G1, eIF4A2, eEF1A1, , Rps12l1and Klhl25). Interestingly, we were also able to identify over 40 genes whose functions have yet to be fully elucidated. Of the genes whose functions have yet to be elucidated over 30 genes are from transcripts that were predicted using *in silico* methods, several of which are for snoRNAs, pseudogenes and one gene that is a predicted protein coding gene (A830018L16Rik). In addition, a number of small, stable RNAs were also identified as being translationally upregulated in TGFβ-induced EMT, which means that they either bind to, or pellet as high-molecular weight species alongside ribosomes. These including, snRNAs (Rnu1a1), snoRNAs

(snora3/65/73a/73b/74b/78, snord12/73a/80/83b/100), the RNA component of RNase P (Rpph1), which is required for tRNA processing, microRNAs (Mir21a/22/24-1/27a/27b/29a/30b/99b/125-b2) and long non-coding RNAs (7S RNA, Malat1).

Recent literature[70,90–92] now indicate that mRNA translation elongation also plays a key role in this fundamental process and the control is far more complex and dynamic than previously anticipated as it is determined by the complex interactions of tRNA, mRNA and elongation factors coupled with dynamics of the ribosome itself. To initiate a successful round of translation elongation Eukaryotic Elongation factors 1A (eEf1a) and ternary complex (TC) containing the amino acylated tRNA and eukaryotic guanine nucleotide exchange factors (GEFs), Eef1b2 and Eef1d, together with the scaffolding protein Eef1g needs to engage on the mRNA transcript and Eef2 subsequently aids in the translocation of tRNAs along the mRNA transcript[93].

As Eef1b2, the main GEF responsible for delivering tRNA was altered on a translation and a localization level during EMT (**Fig. 16b**) we investigated the localization and expression of the remaining components of the TC and only Eef1b2 was altered (**Fig. 16f**).

Subsequently, we determined the level of Eef2 and the inactive, phosphorylated form of Eef2 (pEef2) which displays reduced ribosome binding and has therefore been shown to reduce the elongation process [94,95]. We observed increased expression and confirmed nuclear localization of pEef2 in cells that have undergone TGFβ mediated EMT by co-localization with Smad4 (**Fig. 16f**).

It is known that mTOR signaling regulates the activity of the Eef2 kinase and Eef2 is kept in a predominantly non-phosphorylated state to support efficient

elongation[90]. Thus the increase in inactive pEef2 is in line with the reduced mTOR signaling during EMT presumable through reduced efficiency of translation elongation. Collectively, the expression and localization changes of initiation and elongation factors further substantiate that global and pervasive changes in the translation machinery occur post EMT.

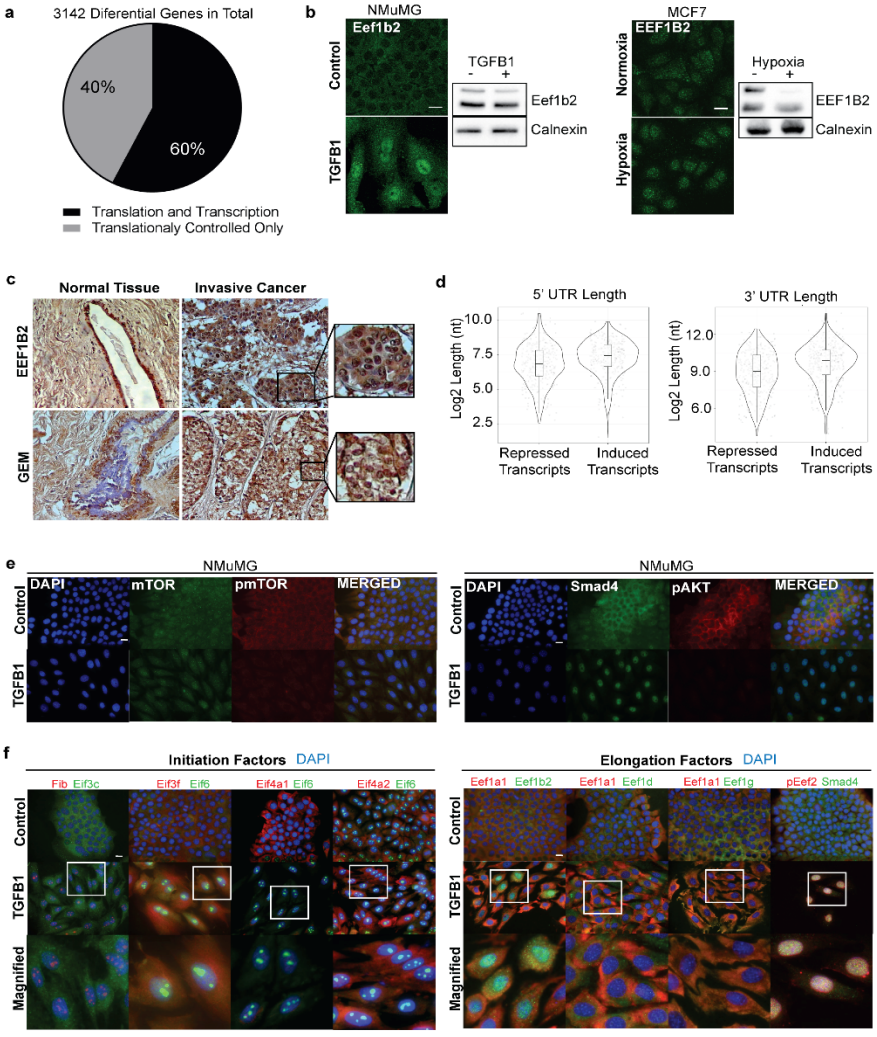

Figure 16. Polymerase profiling demonstrates translational regulation in EMT independent of transcription. (**a**) Pie chart illustrating that approximately 40% of all differential genes are translationally controlled **(b)** Change in both the expression and localization of the protein Eef1b2 in TGFβ driven EMT in NMuMG cells as well as hypoxia (48 hours) driven EMT in MCF7 cells. **(c)** Increased expression and nuclear localization of EEF1B2 and GEM in invasive cancer as compared to normal tissue **(d)** Violin plots illustrating that, on average, the mRNA transcripts of genes that were up-regulated in cells undergoing EMT had longer 5' and 3' UTRs $p < 0.05$**) (e)** Immunofluorescence using phospho-specific antibodies demonstrate the pS2448 on mTOR and pS473 on Akt are lost **(f)** Immunofluorescence of NMuMG cells showing alteration in both expression and localization of initiation factors as well as changes in localization and expression of elongation factors including the phosphorylated form of Eef2. Panel **(a)** was made using analysis from Dr. Matt Parks Immunofluorescence and western blots for panel (**b)** were done by Dr. Brittany Carson. IHC for panel (**c)** was done by Varsha Prakash. Panel (**d)** was produced by Dr. Matthew Parks and modified by me.

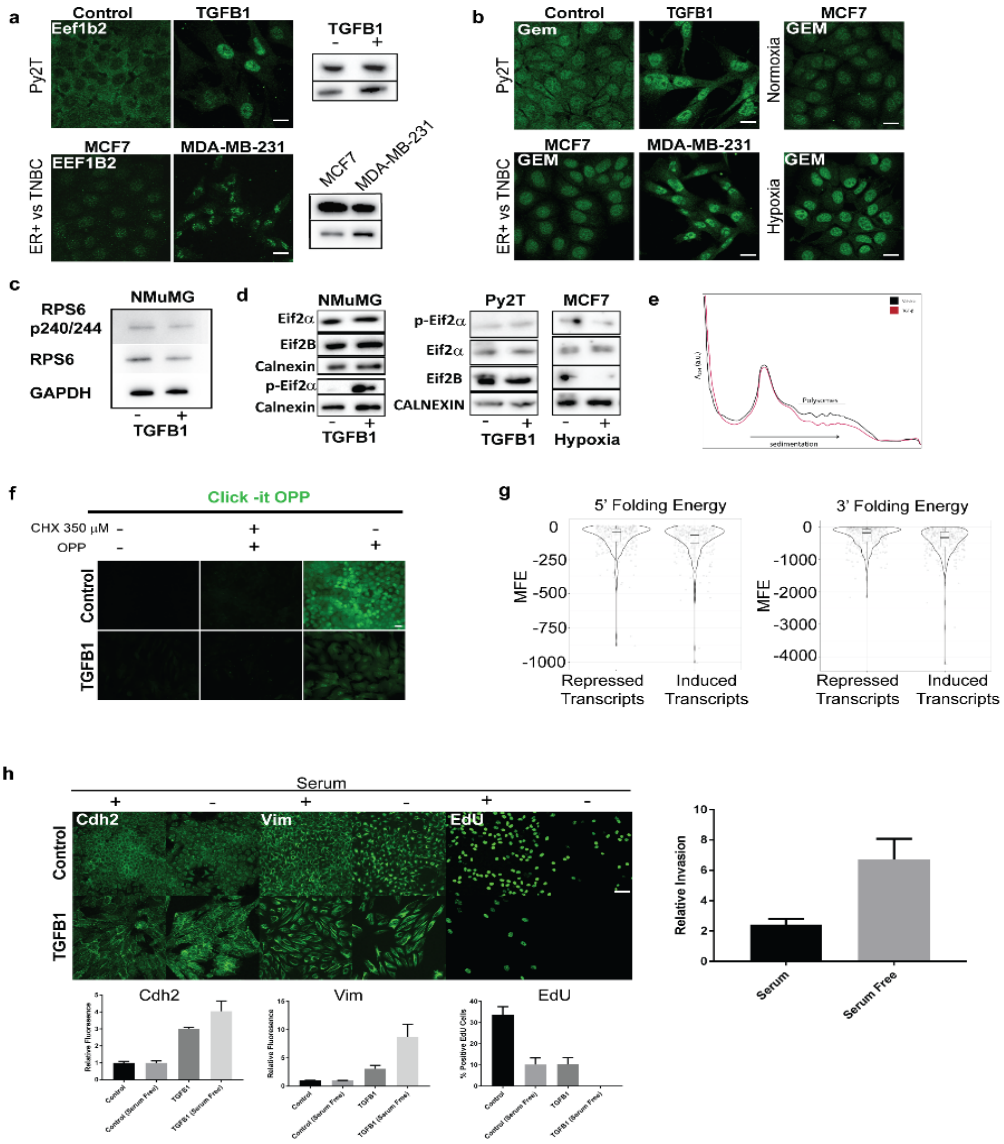

Figure legend and data attributes are on the following page.

Figure S16. Loss of mTOR and gain of pro-mesenchymal mediated translation during EMT. (a) Increase in nuclear localization of Eef1b2 in the TGFβ induced system Py2T as well as the TNBC cell line MDA-MB-231 vs MCF7 human breast cancer cell lines (**b**)Nuclear localization of Gem in the Py2T system as well as the human cell lines MCF7 and MDAMB231 **(c)** Reduction in the expression and phosphorylation of Rps6 at Serine 240/244 in response to TGFβ (**d**) increase in the phosphorylation of the alpha subunit of Eif2 indication a reduction in mTOR driven and cap dependent translation in NMuMG cells. In hypoxic condition with MCF7 cells though the phosphorylation status of the alpha subunit was unchanged there was a significant loss of the GEF EIF2B which has the same overall effect. (**e)** Polysome profiles depicting the active translating pool from either vehicle treated or TGFβ treated cells demonstrating that there is an overall reduction in global translation in response to EMT. (**f**) Global reduction in protein synthesis in response to TGFβ as shown by Click iT OPP fluorescent assay. **(g)** Violin plots comparing the 5' and 3' folding energies of repressed and induced mRNA transcripts in response to TGFβ of genes changed only on a translation level. **(h)** Comparison and quantification of the increase in expression of either Vimentin or N-Cadherin in either a serum starved or serum present state in addition to the proliferative status of the cells as demonstrated by the presence of EdU. Graphs of the relative fluorescence intensity are below representative images of each condition. Quantification of the increase in the invasive capacity of cells that have gone through EMT in a serum starved state as compared to those in a serum added state. Western Blots and immunofluorescence for panels (**a,b,d,h)** were done by Dr. Brittany Carson. Polysome Profiles for panel (**e)** were done by Chad Kurylo. Panel (**g)** was made by Dr. Matthew Parks and modified by me. Invasion assay was performed by Varsha Prakash.

2.1.7 Inhibition of Ribosome Biogenesis and TGFβ Removal Reverses Translational Control

The EMT program is phenotypically inhibited, and can be at least partially reversed, by administration of CX-5461 or by removal of TGFβ (Prakash *et al*). To demonstrate gene expression control of pro-mesenchymal translational program we interfered with *de novo* rRNA biogenesis using the RNA Polymerase I assembly inhibitor, CX-5461[96,97]. NMuMG cells, which had initiated but not completed the EMT program (27 hours of TGFβ treatment), were treated with 100nM CX-5461, a time point and concentration that has no observable impact on DNA replication or apoptotic pathways (REF rDNA). By comparing ribosome profiling and RNASeq data, we detected that CX-5461 treatment reversed the impact of TGFβ for almost 30% of the 1000 genes identified as being translationally controlled (**Fig.17a**). Remarkably, CX-5461 treatment had no significant changes on transcription, suggesting that the observed impacts were solely dependent on the inhibition of ribosome biogenesis.

As expected, the physical characteristics of the mRNA transcripts translationally induced by CX-5461 treatment corresponded to relatively short, unstructured 5'- and 3'-UTRs typified by mTOR sensitive genes; mRNA transcripts that were translationally repressed had the pro-mesenchymal/pro-invasive signature of relatively long, structured 5'- and 3'-UTRs (**Fig. 17b**). To verify that mTOR signaling had reversed in line with the mRNA transcript profiles, we investigated total mTOR, AktpS473 and mTORC1pS2448. As predicted by the mRNA transcripts CX5461 treatment resulted in a small induction mTOR, pAkt and mTORC1 mark pS2448 supporting the notion that mTOR signaling is induced post treatment (**Fig. 17c**). Interestingly, the

observed reduction in the RNA helicase, Eif4a1 post EMT was reversed post CX5461 treatment (**Fig. 17c**). To verify that the pro-mesenchymal program was reduced we investigated Gem and Vimentin expression and as predicted detected reduced expression of both proteins (**Fig. 17d**).

TGF is secreted as an inactive latent complex together with latency-associated proteins (LAPs) where it subsequently binds latent TGFβ-binding protein (LTBP). Both proteins must be liberated from TGFβ for it to be active as a cytokine. Three of the four known members of LTBP protein family (LTBP 1,3 and 4) were identified in our gene set and two of them, LTBP 1 and 3, were observed to reverse translationally efficiency upon CX5461 treatment (**Fig. 17a**). Thrombospondin-1 (Thsb1), an extracellular matrix glycoprotein implicated in TGFβ activation, was also reversed upon CX5461 treatment (**Fig. S17d**). These findings suggest that CX-5461-mediated inhibition of ribosomal biogenesis directly or indirectly interferes with the translationally efficiencies of autocrine loop proteins responsible for TGFβ secretion and activation. Parallel changes in global translational efficiency were observed for 330 of 1100 transcripts found in the CX-5461-impacted gene set when TGFβ was removed from the NMuMG growth media after 48 hours of treatment.

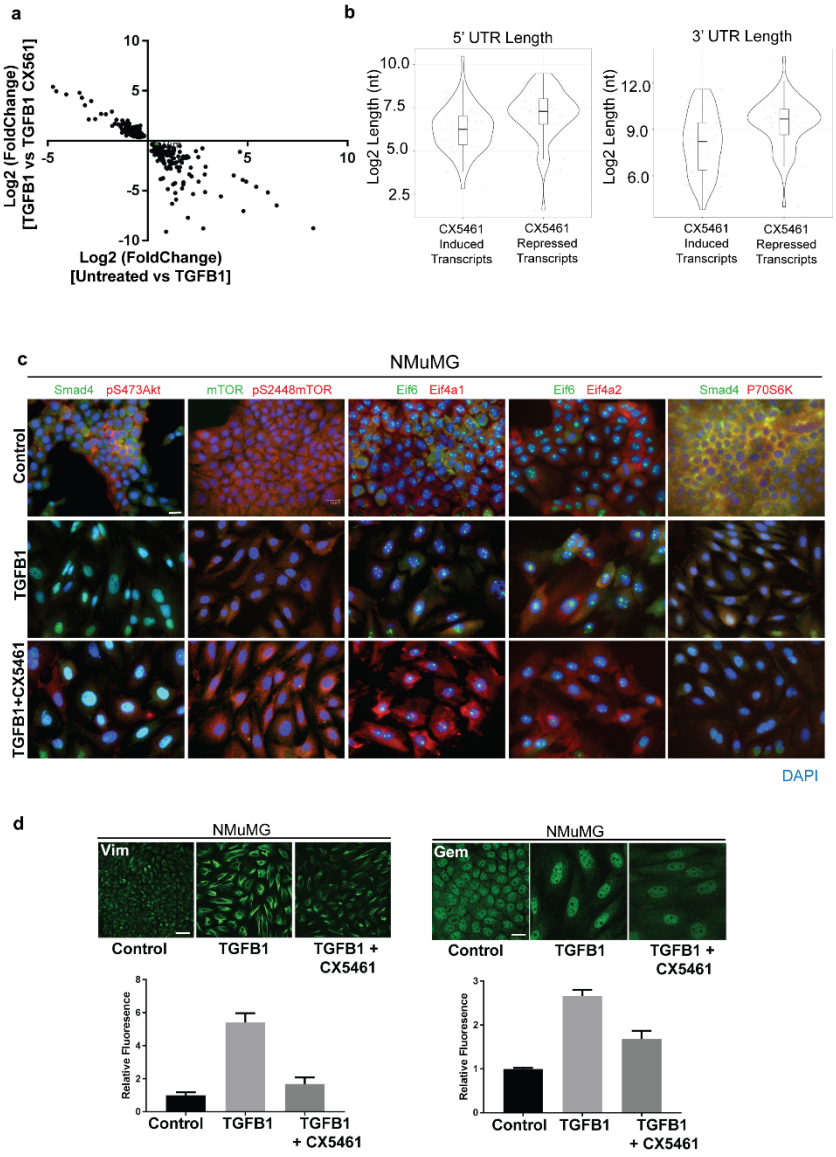

Figure 17. Gene expression is altered by use of a Pol I specific inhibitor. (a) Scatter plot showing the reversal of a subset of genes that were differential only on a translation level in response to the Pol I inhibitor CX5461 (100 nM) **(b)** Violin plots illustrating that the genes that were repressed in response to CX5461 had, on average, longer 5' and 3' UTRs $p<0.05$. Whereas genes induced by CX5461 had, on average, shorter 5' and 3' UTRs $p<0.05$ (**c**) Immunofluorescence using phospho-specific antibodies demonstrating the reduction of phosphorylation at serine 473 and 2448 of Akt and mTOR respectively as well as the loss of expression of Eif4a1. The phospho-marks as well as Eif4a1 expression are regained upon treatment with CX5461 (**d**) Immunofluorescence demonstration the gain in expression of Vim and Gem and subsequent reduction in expression in response to CX561 with the corresponding quantification of the fluorescence signal. Panel **(a)** was made from analysis by Dr. Matt Parks (**b**) was produced by Dr. Matt Parks and modified by me. Immunofluorescence and quantification for Panel **(d)** was done by Dr. Brittany Carson.

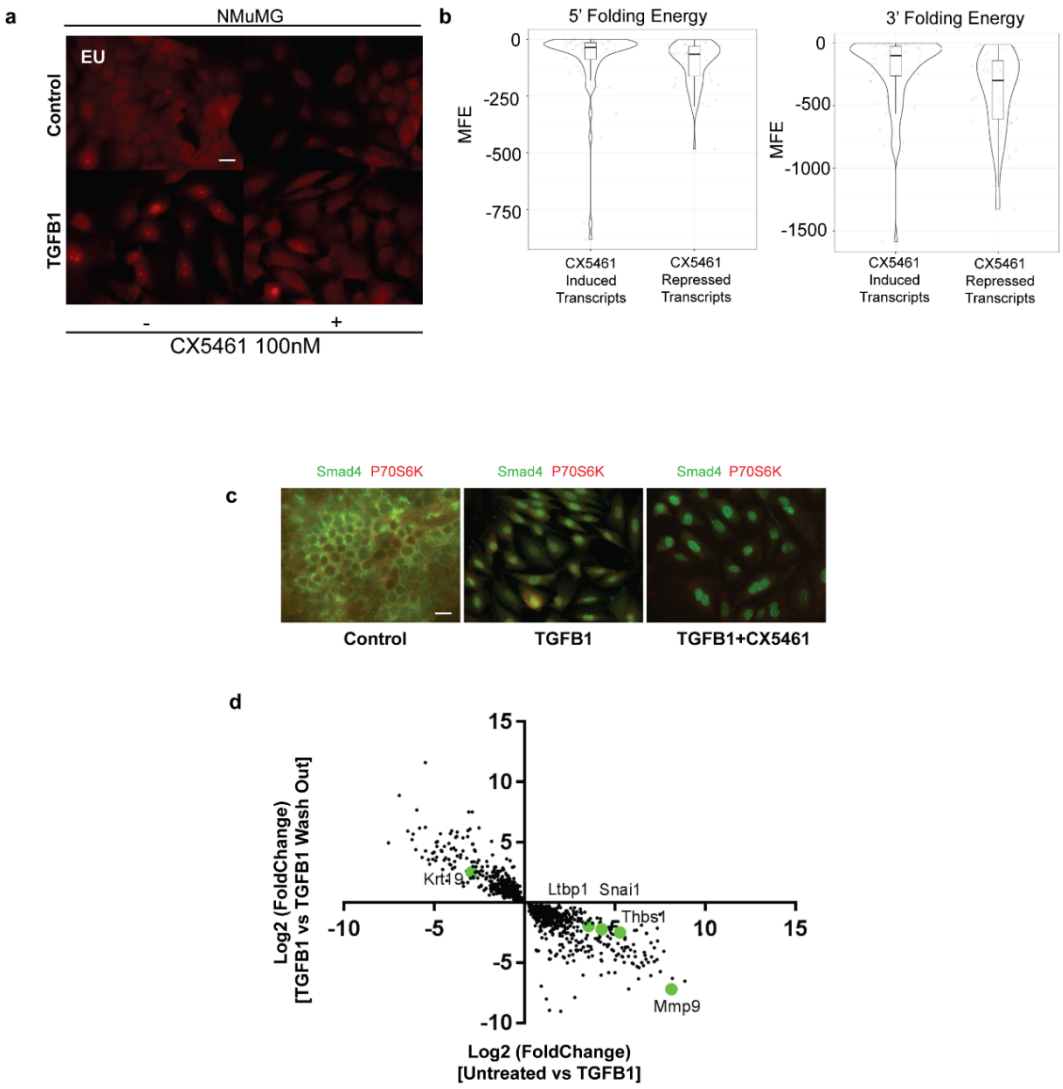

Figure S17. Pol I assembly inhibition can reverse gene expression (**a**) Fluorescent images of EU staining in NMuMG cells both in response to TGFβ and plus and minus CX5461 100nM for 24 hours. **(b)** Folding energy for 5' and 3' UTRs of CX5461 induced or repressed transcripts $p < 0.05$. **(c)** Change in localization of P70S6K to the nucleus in response to TGFβ and change in localization from the nucleus to the cytoplasm with CX5461 treatment for 24 hours. **(d)** Scatter plot of genes that were either induced or repressed at a translation level in response to the removal of TGFβ. Panel **(b)** was made by Dr. Matt Parks and modified by me. Panel **(d)** was made by me using analysis from Dr. Matt Parks.

2.1.8 tRNA Genes are Differentially Utilized During EMT and Inhibition of Ribosome Biogenesis or TGFβ Removal Reverses the tRNA Gene Signature

The tRNA adapter molecule facilitates mRNA decoding during protein synthesis and serves as an integral component of the elongating ribosome by influencing factor interactions and functional ribosome dynamics[98–100]. Specific tRNA species have recently been implicated in regulating cell proliferation and metastasis[101–103]. Given the global changes in mRNA transcripts and the pervasive expression and localization changes of initiation and elongation factors in, we next asked whether additional components of the translation machinery, tRNA was altered to accommodate the EMT switch. Given that the tRNAs retrieved were derived from pelleted ribosomes we refer to this pool of tRNAs as "utilized tRNAs" as they represent a snap-shot of the tRNAs being used by the ribosome in active translation. A bioinformatics strategy was developed where we retrieved amino acid, anti-codon and gene level expression data**.** We observed a modest decrease in the global utilization of Ala, Val, Glu and Lys tRNA isoacceptors and modest increases in Cys, Gln, Ser, Trp, Met, Leu and His tRNA isoacceptors (**Fig. 18a**). The remaining tRNA isoacceptor families exhibited both increased and decreased utilizations for specific isoacceptor sub-types.

Importantly, we could identify and map reads to 197 tRNA genes out of the 260 sequentially distinguishable tRNA genes within the genome, of which, 68 genes (38%) were differentially utilized post TGFβ treatment, resulting in the first identification of a specific subset of tRNA genes that are utilized during EMT (**Fig. 18a**). Interestingly, in our gene set the tRNA gene Glu^{UUC} which has

previously been linked to metastasis was identified to be utilized to a larger extent by the ribosomes post TGFβ treatment[102] (**Fig. 18a**). To identify if this tRNA signature was due to TGFβ signaling and under translation control we analyzed tRNA transcripts utilization with the removal of TGFβ treatment and in the presence of CX5461 (**Fig. 18b**). Of the 68 tRNA genes shown to be differential in response to TGFβ, 27 (41%) were reversible when TGFβ was removed (washout) and 24 (36%) were reversible when CX5461 was used. There was a large overlap where 88% of the tRNA genes were found to be reversible in both conditions suggesting that tRNA utilization by the ribosomes is dependent on the translational changes evoked by the TGFβ signaling. Although we cannot conclusively argue that such changes exclusively reflect changes in the elongating ribosome's composition given that high-molecular weight tRNA synthetase complexes may pellet alongside ribosomes during the profiling procedure, these data minimally suggest that the available TC pool is globally altered during TGFβ-induced EMT.

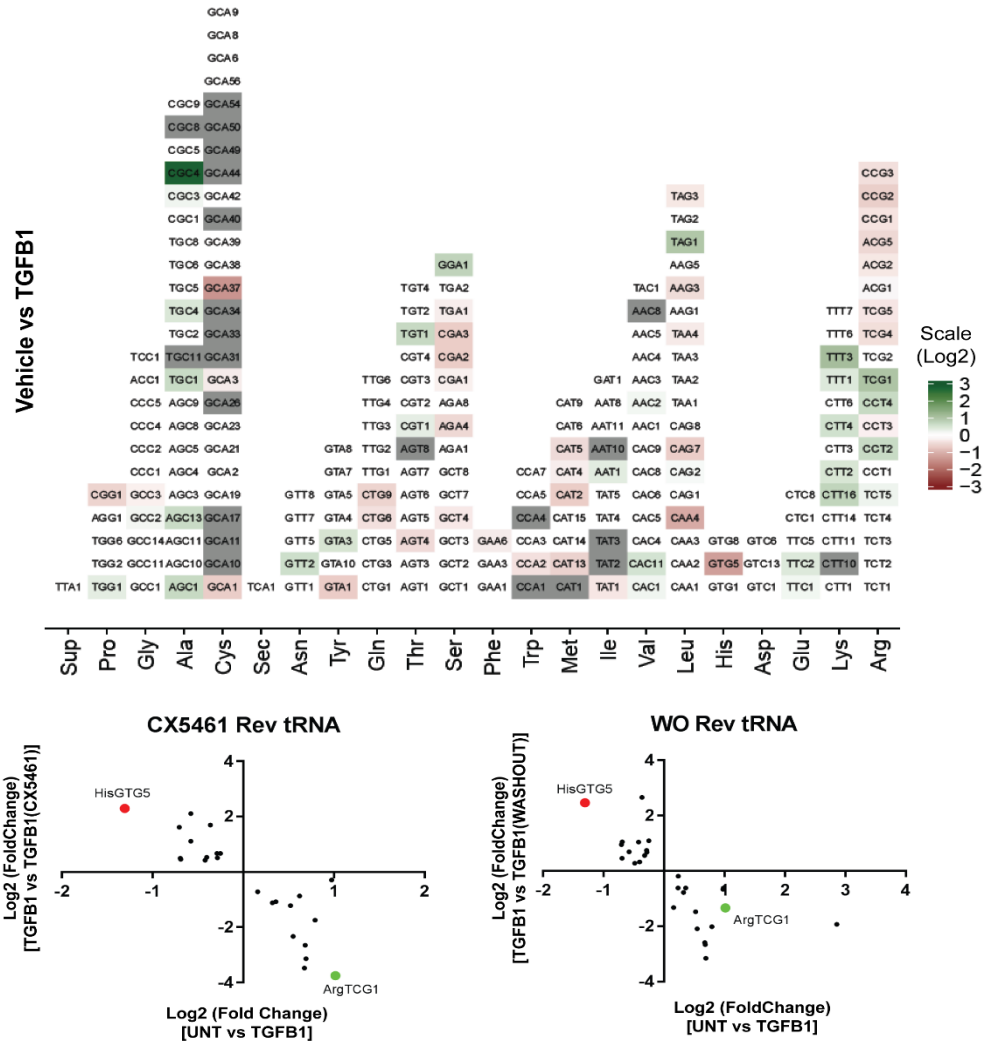

Figure 18. (a) Histogram showing the changes in expression to specific tRNA genes **(b)** Scatter plot of tRNA genes that are reversed either in response to CX5461 100nM for 24 hours or to TGFB removal for 48 hours (WO). The top histogram was produced by Dr. Matthew Parks. The bottom scatter plots were made using analysis by Dr. Matt Parks

2.1.9 rRNA is Altered During EMT and Inhibition of Ribosome Biogenesis and TGFβ removal Reverses the rRNA Changes

The final component and major component of the ribosome is rRNA. Since we did not apply subtractive hybridization of rRNA nor bioinformatically removed rRNA, we retrieved, in an unbiased manner, cleaved rRNA from exposed areas of the ribosome. As the nuclease Rnase I is enzymatically probing solvent

accessible areas of the ribosome and given the preference of Rnase I for single stranded RNA any secondary structure changes, in the rRNA will give rise to differential digestion products in a context dependent manner[33,104,105]. A differential digest could also arise from changes to the protein composition of the ribosomes or differential biding of auxiliary translation factors. Given the observed altered elongation efficiency detected in EMT we specifically asked whether changes in the ribosome in close proximity to the Eef2 binding site could be detected. As the sequence of the rDNA operons in the NMuMG cells are not currently known we instead mapped the retrieved rRNA fragments to the structure of the human 80S ribosome (PDB 3J3A, 3J3B, 3J3D and 3J3F) as the structure of the mouse ribosome has not been determined[106]. A subset of reads mapped to areas in close proximity to the Eef2 binding site and post TGFβ treatment, the retrieved reads were significantly altered 5 $\log_2$ (32 fold) suggesting that enzymatic access of that area was altered (**Fig. 19a**). Collectively this argues that alterations in the binding affinities or frequency of auxiliary factors may manifest itself as areas of differential digests.

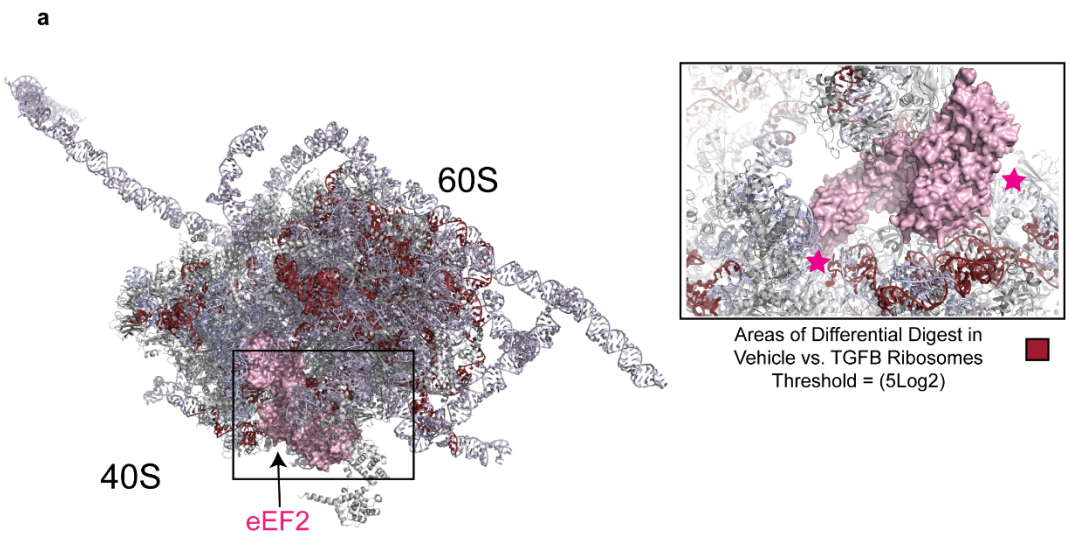

Figure 19. (a) PyMol figure of differentially digested areas of the ribosome. With inset of a zoomed in and rotated view of Eef2 in direct contact with the areas of the ribosome that show large digest changes. The PyMol session utilized analysis by Dr. Matt Parks

2.1.10 Translation Elongation is Altered During EMT such that translation elongation inhibitors selectively target cells in the mesenchymal state

Translation elongation requires the coordination of two key complexes; Given the key role Eef2 plays in the translocation of tRNAs during the elongation phase, we next assessed whether the changed Eef2 would alter the time ribosomes stay in a specific conformational state. It has recently been shown that the distribution of mRNA fragment lengths can be linked to the confirmation state of the ribosome where un-rotated state of the ribosome mimics a post-termination state of the translation cycle, thus implicating a successful round of translation has occured[107].

Based on the pervasive changes to the translation machinery including tRNA, rRNA and specifically elongation factor and TC accessibility to accommodate the EMT switch we tested whether cells transitioning from an epithelial state to a mesenchymal state will respond differently to translation

elongation inhibitors. Accordingly, we treated NMuMG cells in the presence or absence of TGFβ with increasing concentrations of the translation elongation inhibitor Didemnin B. Cells undergoing EMT were much more sensitive to Didemnin B, as we observed little to no cell death in control cells, while we observed pronounced killing of cells that were exposed to TGFβ in line with EMT cells displaying marked differences in translation elongation efficiency (**Fig. 20a,b**).

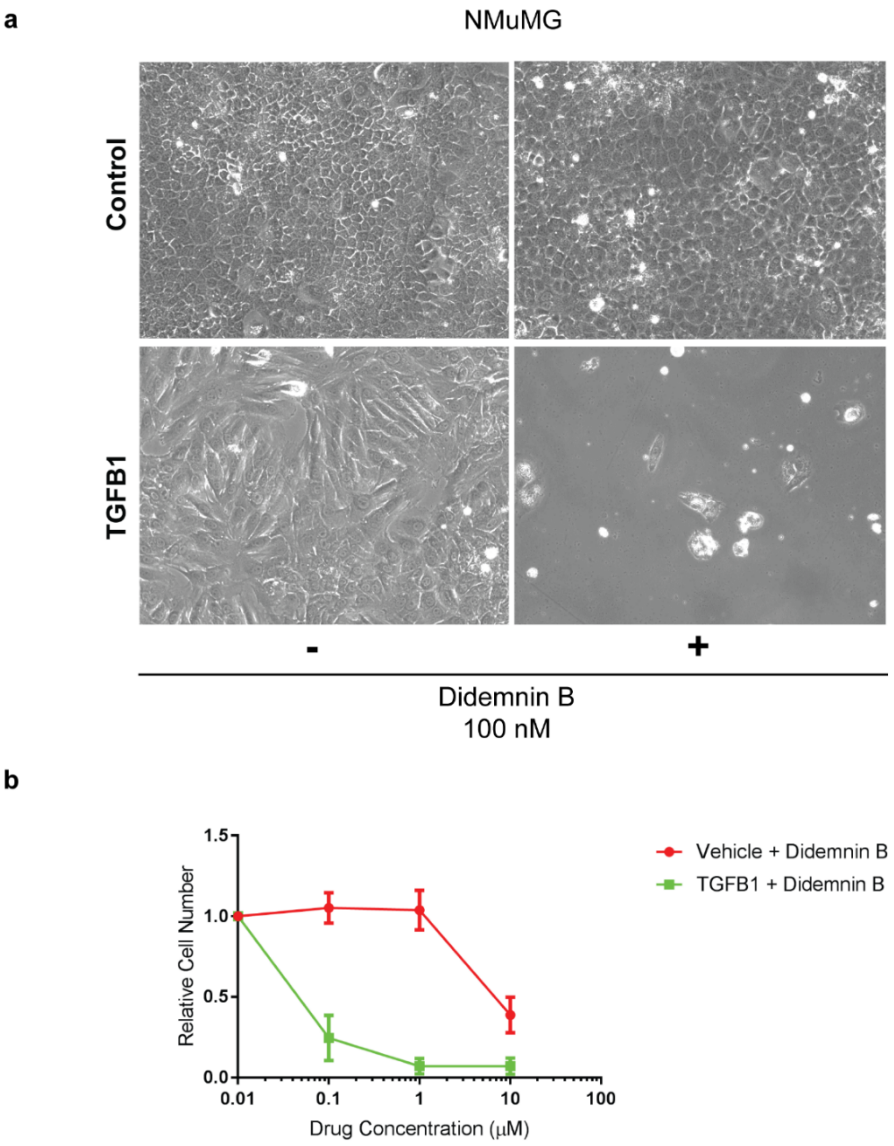

Figure 20. (a) Phase contrast images of vehicle and TGFB1 cells either with addition of 100 nM Didemnin B or DMSO**. (b)** Graph of the quantification of cell number versus Didemnin B concentration showing differing effects on TGFB treated cells.

2.2 DISCUSSION

The regulation of the cell identity switch of EMT has primarily been attributed to transcriptional regulation with few studies indicating a contributing role of translation regulation[18,41,52]. Building upon our previous findings demonstrating an increase in rDNA transcription during EMT in the absence of proliferation, we employed a modified ribosome profiling method and identified that pervasive genome-wide translation regulation occurs during EMT. Importantly, we identified that the translational changes that occurs during EMT involves the translation machinery and components thereof. Three novel EMT markers Eif6, Eef1b2 and Cttnbp2nl were identified and expression changes of these proteins were shown to be generalizable irrespective of EMT stimuli. Importantly, all invasive tumor samples showed increased tissue expression of the newly identified EMT markers compared to normal tissues suggesting that these new markers may have clinical relevance as prognostic indicators of invasive disease in combination or separately[108,109]. Further clinical studies including more human tissue samples will be required to address this in more detail.

By the genome-wide analysis of translation control during EMT, we unexpectedly identified that the mRNA transcripts which were reduced, had shorter and less structured 5' and 3' UTRs, all characteristics of mTOR sensitive genes (**Fig 16d; S16g**)[68,85]. Prior literature has elegantly shown that reduced proliferation is necessary for cells to undergo EMT but mainly on a gene specific transcriptional level. Our data provides a compelling explanation as to how this switch is mediated on a global and translation level by the reduction of mTOR[110]. Building on that observation in the context of an even further reduced mTOR signaling milieu we were able to induce an augmented EMT including increased invasiveness. Reduced mTOR signaling has mostly been associated with an

amino acid starvation state and despite that global translation is reduced (**Fig. S16c, S16f**) we observe that it can be further reduced by serum starvation[45,87]. This argues that during a nutrient limited state, such as within in a rapidly expanding tumor mass, cells are not prevented from undergoing EMT, rather it suggests that during some starvation states i.e. hypoxia, may instead promote EMT. Indeed, previous studies have already demonstrated that hypoxia reduces mTOR signaling and hypoxia has been shown to induce EMT[21]. Our data argues that the link between the two may at least partly, lie in restricting the canonical translation mTOR-dependent proliferative program to promote the non-canonical, pro-mesenchymal migratory translational program. Moreover, it has been reported that cells that transition from an epithelial state to a mesenchymal state can adopt a stem cell like phenotype as well as the need for an inhibition of mTOR in order to induce a paused pluripotent state[22,48,88,89]. This notion is further supported in that we also observe similar EMT markers being changed in our hypoxia induced EMT model (**Fig. S14d**) and that markers of cap-dependent translation are reduced (**Fig. S16d**). The genes translationally induced by TGFβ were instead typified by longer and more structured 5' and 3' UTRs (**Fig. 16d; S16g**) which have been shown to be feature of oncogenes translated by non-canonical, cap-independent translation mechanisms[63]. That these translational programs are remarkable coordinated on a global level to reduce proliferation and promote EMT was shown by inhibition of ribosome biogenesis by CX5461 and TGFβ removal resulted in reversed of these translational programs. Subsequently mTOR program was regained and the pro-mesenchymal program reduced.

It has previously been shown that mTOR signaling activates rDNA transcription in order to promote cell division[45,111]. We therefore conclude that

the observed increase in rDNA transcription must be driven by a distinct and mTOR independent signaling cascade, uncoupling mTOR-dependent rDNA transcription during cell growth from mTOR independent rDNA transcription during EMT. These data also argue that rDNA transcription during EMT regulates the mTOR translation control as perturbation by CX5461 and by extension the making of new ribosomes induces the mTOR sensitive genes and reduces the mTOR insensitive genes without any detectable transcriptional changes.

Recently, it has been recognized that tRNAs play an important role in translation during disease states and aberrations in tRNA expression may be sufficient to induce disease states[103,112,113]. There has been evidence that proliferating cells utilize a different tRNA gene set than differentiating cells thus suggesting that there are distinct tRNA programs for different cellular programs. Our experimental strategy and retrieved data substantiates this observation, as we observe tRNA utilization differences during EMT[101,113]. In line with reduced mTOR signaling we also observe a decrease in the utilization of tRNA genes coding for the amino-acids that induce mTOR signaling such as leucine, arginine and serine. In line with the mRNA transcripts changing, perhaps the most striking observation was the change in the utilized tRNA signature with CX5461. These results directly link the tRNA gene set used for the translation of specific mRNA transcripts and when *de novo* ribosome biogenesis is halted the utilized tRNA gene signature changes. Interaction between the ribosome, tRNA and initiation factors has previously been indicated to be involved in the efficiency by which translation can occur. It could also indicate a feedback mechanism where disruption of Pol I transcription leads to changes in Pol III transcription thus altering the expression of a specific tRNA gene set. It has

previously been shown that there are communication between the Pol I, Pol II and Pol III as Pol III transcribes a physical component of the ribosome (5S rRNA) in addition to tRNAs[114]. One implication of this observation is that the Pol I specific inhibition alters the rRNA pool and correspondingly the tRNAs that are used in executing of the EMT program. These observations are intriguing particularly given the recent findings that specific tRNA gene expression can lead to specific genes being translated more efficiently and directly affecting metastatic outcome[102,115].

It has been widely known within the translation field that the process of elongation is crucial for mRNA translation and that this process is highly dynamic, transient and is directly regulated by the interactions of tRNA, mRNA and elongation factors coupled with dynamics of the ribosome itself. These observations have been elegantly elucidated using cryo-EM and smFRET of ribosomes in various translation elongation stabilized states[116]. Depending on the configuration of the ribosome and which components are actively engaged with the translating ribosome, at a given time, will directly affect the translation elongation dynamics[117]. So far these mechanisms have only been studied in the context of *in-vitro* re-constituted ribosomes where labeled complexes of the ribosomes have been followed in real-time. Given that the mRNA, tRNA and dramatic changes to the components of the elongation machinery including Eef1b2, we interpret that the elongation dynamics of ribosomes differ in EMT cells. Eef1b2 is one component of a hetero-trimer protein where it acts as the main guanosine-nucleotide exchange factor (GEF)[93]. The bacterial homolog of the eukaryotic GEF is Ts and has been shown to aid in the formation and dissociation of TC which comprises of amino-acylated tRNA and an elongation factor Tu, responsible for delivering the appropriate tRNA to actively translating

ribosomes[118]. In bacteria it has been shown that formation of TC is lower in the absence of Ts thus translation can occur albeit with a lower elongation efficiency[119]. Given that the mammalian homolog of Ts, Eef1b2, performs a similar function, changes in the cytoplasmic concentration would affect the speed and/or efficiency of ternary complex formation and therefore slow down the rate of ternary complex formation and by extension the rate of translation. To the best of our knowledge, this study is the first to demonstrate that components that aid in the formation of TC is modulated in response to extracellular signaling in a mammalian cellular system. Most strikingly is the redistribution of the protein from a predominantly cytoplasmic localization to a previously unreported nuclear localization whereas there were no observable changes in the localization of Eef1g or Eef1d (**Fig. 16f**) and that this nuclear expression also is found in the human invasive tissue samples. Although nuclear localization has not previously been reported for Eef1b2, Eef1a has previously been reported to play an important role in regulating the nuclear export of SNAG domain contain proteins such as Snail in EMT[120]. This also opens up the possibility that Eef1b2 might display an as yet unknown function in the nucleus. Further evidence of translation elongation alterations during EMT was evident by the increased expression of the inactive phosphorylated Eef2 and its corresponding localization to the nucleus. In a similar manner to the nuclear localization of Eef1b2, the TGFβ dependent localization of pEef2 raises the question whether it plays an unanticipated role in the nucleus. A decrease in Eef2 activity has been shown to impact the translocation step of translation and our data corroborated this by demonstrating that the conformation of the ribosomes in cells post EMT were predominantly in a rotated state indicating that measurable differences in translation elongation has occurred. More

importantly these alterations in translation elongation rendered cells which were going through or had gone through EMT more sensitive to translation elongation inhibitors and were selectively killed by elongation inhibitors such as Didemnin B leaving the control, untreated cells modestly unaffected.

To gain further insight into the pervasive elongation changes during EMT we also investigated rRNA, the major constituent of the ribosome in areas in close proximity to Eef2. Our modified protocol allowed us to detect differences in the digested rRNA providing evidence that the ribosomes in the cells post EMT might be composed of different proteins or display differential biding of auxiliary translation factors. These digest differences were reversed in a TGFβ and CX5461 dependent manner suggesting that these changes mirror tRNA and mRNA changes. Within the limit of the experimental design we could not specifically elucidate *how* these digest differences arose and specifically what physical constituent of the ribosome or its interacting partners gave rise to the digest patterns. The observations are provocative and further biochemical analysis will be required in order to fully understand the mechanism behind the observed changes in rRNA.

We show that all products of the three known RNA polymerases are altered in the EMT process and more importantly our data strongly argues the existence of two specific translation program that needs to be regulated to execute the EMT program, a reduction in the mTOR dependent and induction of a pro-mesenchymal translation program. By perturbing the making of *de novo* rRNA by CX5461, we believe that the context specific ribosomes generated in EMT will utilize a specific set of tRNAs, elongation factors to translate the necessary pro-mesenchymal mRNAs enabling the cell to switch its identity from an epithelial cell to a mesenchymal state. These findings also highlight one

possible explanation as to why anti-proliferative drugs are ineffective at stopping metastatic spread and outlines an alternative therapeutic strategy whereby the use of translation inhibitors in combination with anti-proliferative drugs may more effectively combat metastatic spread. Further understanding of translation regulation and the mechanisms that underlie those changes are required and warranted. The findings of this research can be summarized in the following model where we hypothesize that there are two distinct translation systems that are coordinated in order to execute the EMT program.

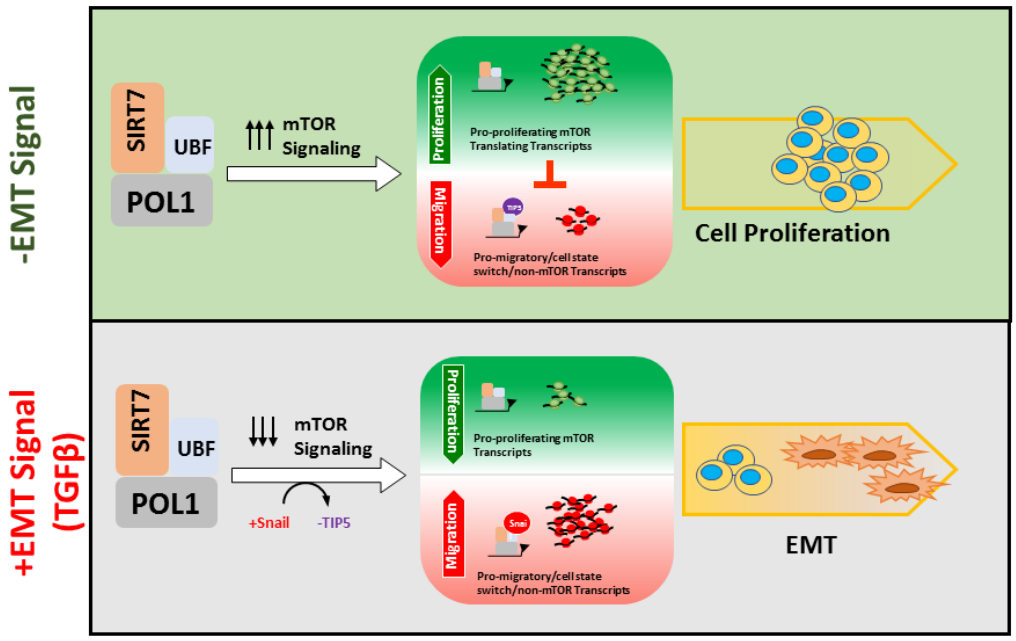

Figure 21. Proposed model of how the proliferative translation program can be reduced in order to allow the execution of the pro-mesenchymal/pro-migratory translation program. The model incorporates findings from Prakash *et al* and Dass *et al*.

CHAPTER 3

CONCLUSION AND FUTURE DIRECTION

In the course of these studies we have endeavored to understand the molecular mechanisms of how tumours develop and subsequently how they are disseminated to other organ systems through metastatic spread. What has come to light is the importance of translation control and perhaps more surprisingly the importance of rDNA transcription. We were able to demonstrate that rDNA transcription plays a key role in facilitating how cells go through EMT and perhaps more surprisingly that changes in rDNA transcription can directly influence gene expression and alter translation programs.

3.1 rDNA TRANSCRIPTION DRIVES PROLIFERATION AND DIFFERENTIATION

The importance of rDNA transcription was highlighted as a functional way Wnt5a exerts its tumour suppressive effect and lead to the observation that there are endogenous signals that are responsible for reducing rDNA transcription in addition to those previously identified as having increased rDNA transcription[121–123]. Given the importance of rDNA transcription to the creation of new ribosomes, the energetic cost of ribosome production and critical role ribosome concentration plays in cell cycle it is not surprising that there are signaling systems in place that tightly regulate rDNA transcription[124–127].

The role of rDNA transcription has not yet been implicated in the process of EMT nor has it directly been linked to the dissemination of cancerous cells. In the manuscript Prakash *et al.* we were able to demonstrate that there is an induction of rDNA transcription in the absence of proliferation. The first time, to the best of our knowledge, that the uncoupling of rDNA transcription occurs as

cells go through EMT in a multitude of systems. There is a strong implication of the uncoupling of rDNA transcription and a cellular identity switch being a general phenomenon that occurs within multiple contexts including both developmental and disease states.

In order to understand what role, if any, the increased rDNA transcription has on the process of EMT we implemented a modified ribosome profiling method. In this case we did not biochemically remove rRNA using subtractive hybridization methods nor did we remove them bioinformatically, in fact, all RNA sequences retrieved were analyzed. Working on the hypothesis that rDNA transcription and by extension the making of new ribosomes altered the translational landscape of cells that are going through EMT. Ribosome profiling would allow for a global snap shot of the translational program, the first such overview, as cells transition through EMT.

3.2 CHANGES TO THE TRANSLATION MACHINERY IN EMT ARE PERVASIVE

Upon implementation and subsequent analysis of the ribosome profiling data we observed the first evidence of translation control that is present as cells go through EMT. Surprisingly, in the NMuMG system, approximately 40% of all genes changes on a translation level were changed *only* on a translation level without the corresponding change in mRNA transcript level. That there was such a large degree of global translation control was unexpected and that it would be as pervasive was not previously documented.

The global changes in translation was not only restricted to transcripts that changes only on a translation level, upon further investigation we observed changes to translation elongation factors, initiation factors, tRNAs, tRNA delivery complexes and aminoacyl synthetases. In fact, there was rarely a part

of the global translation machinery that was unchanged. Perhaps the most surprising observation was the changes in localization of translation related proteins, with a small example of such proteins like Eif4a1, P70S6K and Eif3c going to the nucleus. Such wide-scale trafficking of translation related proteins to the nucleus has not previously been reported in the context of EMT. It is worth of further exploration and study in order to understand the signal that instructs the proteins to go to the nucleus.

3.3 mTOR SIGNALING IS REDUCED AS CELLS GO THROUGH EMT

Perhaps the most striking observation that came from these studies was the reduction of mTOR as cells transition from one cell identity to another. The observation that mTOR signaling is reduced may be a functional cause of the lack of proliferation noted in the Prakash et al manuscript. It also demonstrated that global translation as a whole was reduced as cells made the transition. The observation strongly argues that there are two distinct and independent translation programs in EMT that must be coordinated in order for cells to execute the EMT program.

These observations argue that there is a pro-proliferative mTOR driven translation program that consumes the majority of the cells translation capacity in addition to that there is a pro-mesenchymal/pro-migratory translation that can be executed in the context of a reduced translational background. We were able to identify several pro-mesenchymal genes that are regulated only on a translational level. These results have been recapitulated in multiple systems, particularly in a developmental system like the chick in which there is endogenous levels of multiple signaling pathways.

3.4 GLOBAL TRANSLATION PROGRAMS IN EMT HINGE ON rDNA TRANSCRIPTION

Current thinking has mTOR signaling as a unidirectional master translation regulation pathway. However, in the context of increased rDNA transcription and its correspondent uncoupling during EMT we were able to demonstrate that you can alter mTOR signaling by reducing rDNA transcription. This is the first observation showing that mTOR signaling status could be altered merely by restricting rDNA transcription. The alterations were not merely restricted to the activity status of mTOR and accompanying kinases (Akt) but to mTOR sensitive genes (like ribosomal proteins).

The implication of this observation is that there must be a direct monitoring or feedback loop that allows mTOR to be intimately informed of the status of rDNA transcription. This observation itself is somewhat intuitive given that mTOR can regulate translation in response to nutrient availability. It would make sense that there would be some feedback mechanisms in order to also co-ordinate ribosome biogenesis with nutrient availability and translational need.

3.5 FUTURE DIRECTIONS

In order to truly understand the ramifications of the findings put forth within this thesis, further study into how rDNA transcription is organized and what actually takes place within the nucleus upon rDNA transcription is needed. At this point in time there is a paucity of understanding about the intricacies and ramifications of nucleolar biology, particularly how proteins are trafficked into and out of the nucleolus or potentially how nucleolar shape alters the function. Though there have been some reviews there have yet to be studies linking these observation to functional biological consequences or disease states[128,129]. It is my belief that the shape and size of the nucleolus is dependent on the rate of

rRNA transcription as the rRNA forms a structural component of the nucleolus. This hypothesis can be tested in future work.

In principle given our demonstrated ability to ribosome profile small amounts of material (5 pmols of ribosomes) it may be possible to use cell sorting techniques in order to isolate small amounts of biologically relevant cells to ribosome profile them in order to determine how their translation landscapes are altered. It would also be possible to detect changes in all three RNA polymerases in heterogenous populations like those that occur in tumors.

A stumbling block to determining different rDNA sequences is that the genomic sequence of all rDNA genes for the human and the mouse are not known. This lack of information makes it difficult to get primers or develop tools to explore whether or not there may be different rDNAs (at a sequence level) that are expressed in terms of the different translation programs or in another context dependent manner.

A key finding that also emerged from these studies was that translation elongation inhibitors can selectively target cells that are going through EMT. The translation landscape is sufficiently altered that one can selectively target these cells with translation elongation inhibitors. In addition, there has been *in vivo* work in mice demonstrating that reducing rDNA transcription using the Pol I inhibitor CX5461 can both reduce tumor volume and metastasis. These observations are worthy of follow up studies particularly given the fact that using both proliferation inhibitors as well as translation inhibitors in combination may be of better therapeutic value to patients in treating both the primary tumour and potentially any metastatic spread.

APPENDIX

Reagents needed for NMuMG Growth and Ribosome Profiling

Media composition for NMuMG cells:

500 mL DMEM (High Glucose [4.5 g/L] from Gibco/Invitrogen Cat#11965-018 (12 bottles))

10% (50 mL) Fetal Bovine Serum (Atlanta Biologicals Cat# S11150 Lot# B1040)

1% (5mL) Penicillin/Streptomycin (Gibco/Invitrogen Cat# 15140)

Insulin @ final concentration of 10 µg/mL (500µL of a 10mg/mL insulin solution from bovine pancreas (Sigma Cat#I0516-5ML))

Reagent List

DMEM (High Glucose [4.5 g/L] from Gibco/Invitrogen Cat#11965-018

Atlanta Biologicals Cat# S11150 Lot # B1040

Penicillin/Streptomycin Gibco/Invitrogen Cat# 15140

10 mg/mL insulin solution from bovine pancreas (Sigma Cat# I0516-5ML)

Recombinant Human TGF-beta-1, CF (R&D systems 240-B-002/CF)

10 cm tissue culture plates Corning Cat# 430167

15 cm tissue culture plates Corning Cat# 430599

T225 Tissue Culture Flasks BD Cat# 353138

(Falcon vented cap sterile tissue culture flasks)

Cycloheximide EMD Millipore 239764-1GM

TGFβ1 Re-suspension:

TGFβ1 should be re-suspended using a 0.1% BSA/ 4mM HCl buffer. I have a stock solution of the re-suspension buffer in the 4 °C fridge in cell culture. Stock solution of TGFβ1 should be 10 ng/µL.

PCR Primers for Library Indexing:

Forward PCR primer:

5′-AATGATACGGCGACCACCGAGATCTACAC-3′

ACGACT

CAAGCAGAAGACGGCATACGAGATAGTCGTGTGACTGGAGTTCAGACGTGTGCTCTTCCG

ATCAGT

CAAGCAGAAGACGGCATACGAGATACTGATGTGACTGGAGTTCAGACGTGTGCTCTTCCG

CAGCAT

CAAGCAGAAGACGGCATACGAGATATGCTGGTGACTGGAGTTCAGACGTGTGCTCTTCCG

CGACGT

CAAGCAGAAGACGGCATACGAGATACGTCGGTGACTGGAGTTCAGACGTGTGCTCTTCCG

GCAGCT

CAAGCAGAAGACGGCATACGAGATAGCTGCGTGACTGGAGTTCAGACGTGTGCTCTTCCG

TACGAT

CAAGCAGAAGACGGCATACGAGATATCGTAGTGACTGGAGTTCAGACGTGTGCTCTTCCG

CTGACG

CAAGCAGAAGACGGCATACGAGATCGTCAGGTGACTGGAGTTCAGACGTGTGCTCTTCCG

GCTACG

CAAGCAGAAGACGGCATACGAGATCGTAGCGTGACTGGAGTTCAGACGTGTGCTCTTCCG

Size Marker Sequences:

NI-NI-42

5'-AUGUACACGGAGUCGAGCUCAACCCGCAACGCGACAACGCGA-(Phos)-3'

NI-NI-19

5'-AUGUACACGGAGUCGAGCUCAACCCGCAACGCGA-(Phos)-3'.

NI-NI-20

5'-AUGUACACGGAGUCGACCCAACGCGA-(Phos)-3'

NI-NI-18

5'-AUGUACACGGAGUCGACC-(Phos)-3'

Cloning Linker:

IDT Cloning Linker 1

/5rApp/CTGTAGGCACCATCAAT/3ddC

BIBLIOGRAPHY

1. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144,** 646–674 (2011).

2. Weinberg, R. A. Coming full circle-from endless complexity to simplicity and back again. *Cell* **157,** 267–271 (2014).

3. Gupta, G. P. & Massagué, J. Cancer metastasis: building a framework. *Cell* **127,** 679–695 (2006).

4. Massagué, J. & Obenauf, A. C. Metastatic colonization by circulating tumour cells. *Nature* **529,** 298–306 (2016).

5. Hoshino, A. *et al.* Tumour exosome integrins determine organotropic metastasis. *Nature* **527,** 329–335 (2015).

6. Costa-Silva, B. *et al.* Pancreatic cancer exosomes initiate pre-metastatic niche formation in the liver. *Nat Cell Biol* **17,** 816–826 (2015).

7. Nusse, R. & Varmus, H. E. Many tumors induced by the mouse mammary tumor virus contain a provirus integrated in the same region of the host genome. *Cell* **31,** 99–109 (1982).

8. Nusse, R. & Varmus, H. Three decades of Wnts: a personal perspective on how a scientific field developed. *EMBO J* **31,** 2670–2684 (2012).

9. Wood, L. D. *et al.* The genomic landscapes of human breast and colorectal cancers. *Science* **318,** 1108–1113 (2007).

10. Wan, L., Pantel, K. & Kang, Y. Tumor metastasis: moving new biological insights into the clinic. *Nat Med* **19,** 1450–1464 (2013).

11. Vincent, C. T., Dass, R. A. & Thompson, C. B. A dialogue with Dr. Craig B. Thompson about metabolism and its relevance for tumor growth, progression and metastasis. *Semin Cancer Biol* **22,** 484–488 (2012).

12. Bard, J. B. & Hay, E. D. The behavior of fibroblasts from the developing avian cornea. Morphology and movement in situ and in vitro. *J Cell Biol* **67,** 400–418 (1975).

13. Greenburg, G. & Hay, E. D. Epithelia suspended in collagen gels can lose polarity and express characteristics of migrating mesenchymal cells. *J Cell Biol* **95,** 333–339 (1982).

14. Hay, E. D. & Fischman, D. A. Origin of the blastema in regenerating limbs of the newt Triturus viridescens. *Dev Biol* **3,** 26–59 (1961).

15. Ocaña, O. H. *et al.* Metastatic colonization requires the repression of the epithelial-mesenchymal transition inducer Prrx1. *Cancer Cell* **22,** 709–724 (2012).

16. Nieto, M. A., Huang, R. Y.-J., Jackson, R. A. & Thiery, J. P. EMT: 2016. *Cell* **166,** 21–45 (2016).

17. Vincent, T. *et al.* A SNAIL1-SMAD3/4 transcriptional repressor complex promotes TGF-beta mediated epithelial-mesenchymal transition. *Nat Cell Biol* **11,** 943–950 (2009).

18. LaGamba, D., Nawshad, A. & Hay, E. D. Microarray analysis of gene expression during epithelial-mesenchymal transformation. *Dev Dyn* **234,** 132–142 (2005).

19. Valcourt, U., Kowanetz, M., Niimi, H., Heldin, C.-H. & Moustakas, A. TGF-beta and the Smad signaling pathway support transcriptomic reprogramming during epithelial-mesenchymal cell transition. *Mol Biol Cell* **16,** 1987–2002 (2005).

20. Chen, J., Imanaka, N., Chen, J. & Griffin, J. D. Hypoxia potentiates Notch signaling in breast cancer leading to decreased E-cadherin expression and increased cell migration and invasion. *Br J Cancer* **102,** 351–360 (2010).

21. Sahlgren, C., Gustafsson, M. V., Jin, S., Poellinger, L. & Lendahl, U. Notch signaling mediates hypoxia-induced tumor cell migration and invasion. *Proc Natl Acad Sci U S A* **105,** 6392–6397 (2008).

22. Scheel, C. *et al.* Paracrine and autocrine signals induce and maintain mesenchymal and stem cell states in the breast. *Cell* **145,** 926–940 (2011).

23. Gal, A. *et al.* Sustained TGF beta exposure suppresses Smad and non-Smad signalling in mammary epithelial cells, leading to EMT and inhibition of growth arrest and apoptosis. *Oncogene* **27,** 1218–1230 (2008).

24. Gupta, S. & Maitra, A. EMT: matter of life or death? *Cell* **164,** 840–842 (2016).

25. Nieto, M. A., Sargent, M. G., Wilkinson, D. G. & Cooke, J. Control of cell behavior during vertebrate development by Slug, a zinc finger gene. *Science* **264,** 835–839 (1994).

26. Thiery, J. P., Acloque, H., Huang, R. Y. J. & Nieto, M. A. Epithelial-mesenchymal transitions in development and disease. *Cell* **139,** 871–890 (2009).

27. Hazan, R. B., Kang, L., Roe, S., Borgen, P. I. & Rimm, D. L. Vinculin is associated with the E-cadherin adhesion complex. *J Biol Chem* **272,** 32448–32453 (1997).

28. Miettinen, P. J., Ebner, R., Lopez, A. R. & Derynck, R. TGF-beta induced transdifferentiation of mammary epithelial cells to mesenchymal cells: involvement of type I receptors. *J Cell Biol* **127,** 2021–2036 (1994).

29. Moustakas, A. & Heldin, C.-H. Mechanisms of TGFβ-Induced Epithelial-Mesenchymal Transition. *J Clin Med* **5,** (2016).

30. Evdokimova, V. *et al.* Translational activation of snail1 and other developmentally regulated transcription factors by YB-1 promotes an epithelial-mesenchymal transition. *Cancer Cell* **15,** 402–415 (2009).

31. Sheridan, C. M. *et al.* YB-1 and MTA1 protein levels and not DNA or mRNA alterations predict for prostate cancer recurrence. *Oncotarget* **6,** 7470–7480 (2015).

32. Ingolia, N. T., Ghaemmaghami, S., Newman, J. R. S. & Weissman, J. S. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* **324,** 218–223 (2009).

33. Ingolia, N. T., Brar, G. A., Rouskin, S., McGeachy, A. M. & Weissman, J. S. The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. *Nat Protoc* **7,** 1534–1550 (2012).

34. Archer, S. K., Shirokikh, N. E., Beilharz, T. H. & Preiss, T. Dynamics of ribosome scanning and recycling revealed by translation complex profiling. *Nature* **535,** 570–574 (2016).

35. Ingolia, N. T. *et al.* Ribosome profiling reveals pervasive translation outside of annotated protein-coding genes. *Cell Rep* **8,** 1365–1379 (2014).

36. Calviello, L. *et al.* Detecting actively translated open reading frames in ribosome profiling data. *Nat Methods* **13,** 165–170 (2016).

37. Ji, Z., Song, R., Regev, A. & Struhl, K. Many lncRNAs, 5'UTRs, and pseudogenes are translated and some are likely to express functional proteins. *elife* **4,** e08890 (2015).

38. Stern-Ginossar, N. *et al.* Decoding human cytomegalovirus. *Science* **338,** 1088–1093 (2012).

39. Blobel, G. & Sabatini, D. Dissociation of mammalian polyribosomes into subunits by puromycin. *Proc Natl Acad Sci U S A* **68,** 390–394 (1971).

40. Salunkhe, M., Wu, T. & Letsinger, R. L. Control of folding and binding of oligonucleotides by use of a nonnucleotide linker. *J Am Chem Soc* **114,** 8768–8772 (1992).

41. Kalluri, R. & Weinberg, R. A. The basics of epithelial-mesenchymal transition. *J Clin Invest* **119,** 1420–1428 (2009).

42. Iser, I. C., Pereira, M. B., Lenz, G. & Wink, M. R. The Epithelial-to-Mesenchymal Transition-Like Process in Glioblastoma: An Updated Systematic Review and In Silico Investigation. *Med Res Rev* **37,** 271–313 (2017).

43. Piek, E., Moustakas, A., Kurisaki, A., Heldin, C. H. & Dijke, P. ten. TGF-(beta) type I receptor/ALK-5 and Smad proteins mediate epithelial to mesenchymal transdifferentiation in NMuMG breast epithelial cells. *J Cell Sci* **112 (Pt 24),** 4557–4568 (1999).

44. Donati, G., Montanaro, L. & Derenzini, M. Ribosome biogenesis and control of cell proliferation: p53 is not alone. *Cancer Res* **72,** 1602–1607 (2012).

45. Iadevaia, V., Huo, Y., Zhang, Z., Foster, L. J. & Proud, C. G. Roles of the mammalian target of rapamycin, mTOR, in controlling ribosome biogenesis and protein synthesis. *Biochem Soc Trans* **40,** 168–172 (2012).

46. Topisirovic, I. & Sonenberg, N. mRNA translation and energy metabolism in cancer: the role of the MAPK and mTORC1 pathways. *Cold Spring Harb Symp Quant Biol* **76,** 355–367 (2011).

47. White, R. J. RNA polymerases I and III, growth control and cancer. *Nat Rev Mol Cell Biol* **6,** 69–78 (2005).

48. Ye, X. & Weinberg, R. A. Epithelial-Mesenchymal Plasticity: A Central Regulator of Cancer Progression. *Trends Cell Biol* **25,** 675–686 (2015).

49. Tsai, J. H., Donaher, J. L., Murphy, D. A., Chau, S. & Yang, J. Spatiotemporal regulation of epithelial-mesenchymal transition is essential for squamous cell carcinoma metastasis. *Cancer Cell* **22,** 725–736 (2012).

50. Chambers, A. F., Groom, A. C. & MacDonald, I. C. Dissemination and growth of cancer cells in metastatic sites. *Nat Rev Cancer* **2,** 563–572 (2002).

51. Topcul, M. & Cetin, I. Clinical significance of epithelial-mesenchymal transition and cancer stem cells. *J BUON* **21,** 312–319 (2016).

52. Moustakas, A. & Heldin, C.-H. Signaling networks guiding epithelial-mesenchymal transitions during embryogenesis and cancer progression. *Cancer Sci* **98,** 1512–1520 (2007).

53. Robichaud, N. *et al.* Phosphorylation of eIF4E promotes EMT and metastasis via translational control of SNAIL and MMP-3. *Oncogene* **34,** 2032–2042 (2015).

54. Chen, Y.-K., Chen, C.-Y., Hu, H.-T. & Hsueh, Y.-P. CTTNBP2, but not CTTNBP2NL, regulates dendritic spinogenesis and synaptic distribution of the striatin-PP2A complex. *Mol Biol Cell* **23,** 4383–4392 (2012).

55. Klinge, S., Voigts-Hoffmann, F., Leibundgut, M., Arpagaus, S. & Ban, N. Crystal structure of the eukaryotic 60S ribosomal subunit in complex with initiation factor 6. *Science* **334,** 941–948 (2011).

56. Ceci, M. *et al.* Release of eIF6 (p27BBP) from the 60S subunit allows 80S ribosome assembly. *Nature* **426,** 579–584 (2003).

57. Gartmann, M. *et al.* Mechanism of eIF6-mediated inhibition of ribosomal subunit joining. *J Biol Chem* **285,** 14848–14851 (2010).

58. Waldmeier, L., Meyer-Schaller, N., Diepenbruck, M. & Christofori, G. Py2T murine breast cancer cells, a versatile model of TGFβ-induced EMT in vitro and in vivo. *PLoS ONE* **7,** e48651 (2012).

59. Lundgren, K., Nordenskjöld, B. & Landberg, G. Hypoxia, Snail and incomplete epithelial-mesenchymal transition in breast cancer. *Br J Cancer* **101,** 1769–1781 (2009).

60. Attar-Schneider, O., Drucker, L. & Gottfried, M. Migration and epithelial-to-mesenchymal transition of lung cancer can be targeted via translation initiation factors eIF4E and eIF4GI. *Lab Invest* **96,** 1004–1015 (2016).

61. Baggiolini, A. *et al.* Premigratory and migratory neural crest cells are multipotent in vivo. *Cell Stem Cell* **16,** 314–322 (2015).

62. Gingras, A. C., Raught, B. & Sonenberg, N. eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation. *Annu Rev Biochem* **68,** 913–963 (1999).

63. Walters, B. & Thompson, S. R. Cap-Independent Translational Control of Carcinogenesis. *Front Oncol* **6,** 128 (2016).

64. Ali, M. U., Ur Rahman, M. S., Jia, Z. & Jiang, C. Eukaryotic translation initiation factors and cancer. *Tumour Biol* **39,** 1010428317709805 (2017).

65. Bhat, M. *et al.* Targeting the translation machinery in cancer. *Nat Rev Drug Discov* **14,** 261–278 (2015).

66. Sokabe, M. & Fraser, C. S. A helicase-independent activity of eIF4A in promoting mRNA recruitment to the human ribosome. *Proc Natl Acad Sci U S A* **114,** 6304–6309 (2017).

67. Wolfe, A. L. *et al.* RNA G-quadruplexes cause eIF4A-dependent oncogene translation in cancer. *Nature* **513,** 65–70 (2014).

68. Hsieh, A. C. *et al.* The translational landscape of mTOR signalling steers cancer initiation and metastasis. *Nature* **485,** 55–61 (2012).

69. Wen, F. *et al.* The tumor suppressive role of eIF3f and its function in translation inhibition and rRNA degradation. *PLoS ONE* **7,** e34194 (2012).

70. Robert, F. *et al.* Altering chemosensitivity by modulating translation elongation. *PLoS ONE* **4,** e5428 (2009).

71. Hinnebusch, A. G., Ivanov, I. P. & Sonenberg, N. Translational control by 5'-

untranslated regions of eukaryotic mRNAs. *Science* **352,** 1413–1416 (2016).

72. Lee, A. S., Kranzusch, P. J., Doudna, J. A. & Cate, J. H. D. eIF3d is an mRNA cap-binding protein that is required for specialized translation initiation. *Nature* **536,** 96–99 (2016).

73. Lee, A. S. Y., Kranzusch, P. J. & Cate, J. H. D. eIF3 targets cell-proliferation messenger RNAs for translational activation or repression. *Nature* **522,** 111–114 (2015).

74. Lee, J.-Y., Kim, H.-J., Rho, S. B. & Lee, S.-H. eIF3f reduces tumor growth by directly interrupting clusterin with anti-apoptotic property in cancer cells. *Oncotarget* **7,** 18541–18557 (2016).

75. Williams-Hill, D. M., Duncan, R. F., Nielsen, P. J. & Tahara, S. M. Differential expression of the murine eukaryotic translation initiation factor isogenes eIF4A(I) and eIF4A(II) is dependent upon cellular growth status. *Arch Biochem Biophys* **338,** 111–120 (1997).

76. Johnson, A. G., Grosely, R., Petrov, A. N. & Puglisi, J. D. Dynamics of IRES-mediated translation. *Philos Trans R Soc Lond, B, Biol Sci* **372,** (2017).

77. Otto, G. A. & Puglisi, J. D. The pathway of HCV IRES-mediated translation initiation. *Cell* **119,** 369–380 (2004).

78. Lukavsky, P. J., Kim, I., Otto, G. A. & Puglisi, J. D. Structure of HCV IRES domain II determined by NMR. *Nat Struct Biol* **10,** 1033–1038 (2003).

79. Kim, I., Lukavsky, P. J. & Puglisi, J. D. NMR study of 100 kDa HCV IRES RNA using segmental isotope labeling. *J Am Chem Soc* **124,** 9338–9339 (2002).

80. Ingolia, N. T., Lareau, L. F. & Weissman, J. S. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell* **147,** 789–802 (2011).

81. Sendoel, A. *et al.* Translation from unconventional 5' start sites drives tumour initiation. *Nature* **541,** 494–499 (2017).

82. Liberman, N. *et al.* DAP5 associates with eIF2β and eIF4AI to promote Internal Ribosome Entry Site driven translation. *Nucleic Acids Res* **43,** 3764–3775 (2015).

83. Sasikumar, A. N., Perez, W. B. & Kinzy, T. G. The many roles of the eukaryotic elongation factor 1 complex. *Wiley Interdiscip Rev RNA* **3,** 543–555 (2012).

84. Cans, C. *et al.* Translationally controlled tumor protein acts as a guanine nucleotide dissociation inhibitor on the translation elongation factor eEF1A. *Proc Natl Acad Sci U S A* **100,** 13892–13897 (2003).

85. Stumpf, C. R., Moreno, M. V., Olshen, A. B., Taylor, B. S. & Ruggero, D. The translational landscape of the mammalian cell cycle. *Mol Cell* **52,** 574–582 (2013).

86. Hay, N. & Sonenberg, N. Upstream and downstream of mTOR. *Genes Dev* **18,** 1926–1945 (2004).

87. Laplante, M. & Sabatini, D. M. mTOR signaling in growth control and disease. *Cell* **149,** 274–293 (2012).

88. Bulut-Karslioglu, A. *et al.* Inhibition of mTOR induces a paused pluripotent state. *Nature* **540,** 119–123 (2016).

89. Scheel, C. & Weinberg, R. A. Cancer stem cells and epithelial-mesenchymal transition: concepts and molecular links. *Semin Cancer Biol* **22,** 396–403 (2012).

90. Gismondi, A. *et al.* Ribosomal stress activates eEF2K-eEF2 pathway causing translation elongation inhibition and recruitment of terminal oligopyrimidine (TOP) mRNAs on polysomes. *Nucleic Acids Res* **42,** 12668–12680 (2014).

91. Yu, C.-H. *et al.* Codon Usage Influences the Local Rate of Translation Elongation to Regulate Co-translational Protein Folding. *Mol Cell* **59,** 744–754 (2015).

92. Uchida, S., Sato, H., Yoneda, M. & Kai, C. Eukaryotic elongation factor 1-beta interacts with the 5' untranslated region of the M gene of Nipah virus to promote mRNA translation. *Arch Virol* **161,** 2361–2368 (2016).

93. Le Sourd, F. *et al.* eEF1B: At the dawn of the 21st century. *Biochim Biophys Acta* **1759,** 13–31 (2006).

94. Heise, C. *et al.* Elongation factor-2 phosphorylation in dendrites and the regulation of dendritic mRNA translation in neurons. *Front Cell Neurosci* **8,** 35 (2014).

95. Carlberg, U., Nilsson, A. & Nygård, O. Functional properties of phosphorylated elongation factor 2. *Eur J Biochem* **191,** 639–645 (1990).

96. Drygin, D. *et al.* Targeting RNA polymerase I with an oral small molecule CX-5461 inhibits ribosomal RNA synthesis and solid tumor growth. *Cancer Res* **71,** 1418–1430 (2011).

97. Quin, J. *et al.* Inhibition of RNA polymerase I transcription initiation by CX-5461 activates non-canonical ATM/ATR signaling. *Oncotarget* **7,** 49800–49818 (2016).

98. Nedialkova, D. D. & Leidel, S. A. Optimization of Codon Translation Rates via tRNA Modifications Maintains Proteome Integrity. *Cell* **161,** 1606–1618 (2015).

99. Beilsten-Edmands, V. *et al.* eIF2 interactions with initiator tRNA and eIF2B are regulated by post-translational modifications and conformational dynamics. *Cell discovery* **1,** 15020 (2015).

100. Johansson, M., Chen, J., Tsai, A., Kornberg, G. & Puglisi, J. D. Sequence-dependent elongation dynamics on macrolide-bound ribosomes. *Cell Rep* **7,** 1534–1546 (2014).

101. Gingold, H. *et al.* A dual program for translation regulation in cellular proliferation and differentiation. *Cell* **158,** 1281–1292 (2014).

102. Goodarzi, H. *et al.* Modulated Expression of Specific tRNAs Drives Gene Expression and Cancer Progression. *Cell* **165,** 1416–1427 (2016).

103. Goodarzi, H. *et al.* Endogenous tRNA-Derived Fragments Suppress Breast Cancer Progression via YBX1 Displacement. *Cell* **161,** 790–802 (2015).

104. Gerashchenko, M. V. & Gladyshev, V. N. Ribonuclease selection for ribosome profiling. *Nucleic Acids Res* **45,** e6 (2017).

105. Spahr, P. F. & Hollingworth, B. R. Purification and Mechanism of Action of Ribonuclease from Escherichia coli Ribosomes. *Journal of Biological Chemistry* (1961). at <http://www.jbc.org/content/236/3/823.citation>

106. Anger, A. M. *et al.* Structures of the human and Drosophila 80S ribosome. *Nature* **497,** 80–85 (2013).

107. Lareau, L. F., Hite, D. H., Hogan, G. J. & Brown, P. O. Distinct stages of the translation elongation cycle revealed by sequencing ribosome-protected mRNA fragments. *elife* **3,** e01257 (2014).

108. Bullwinkel, J. *et al.* Ki-67 protein is associated with ribosomal RNA transcription in quiescent and proliferating cells. *J Cell Physiol* **206,** 624–635 (2006).

109. Carey, L., Winer, E., Viale, G., Cameron, D. & Gianni, L. Triple-negative breast cancer: disease entity or title of convenience? *Nat Rev Clin Oncol* **7,** 683–692 (2010).

110. Vega, S. *et al.* Snail blocks the cell cycle and confers resistance to cell death. *Genes Dev* **18,** 1131–1143 (2004).

111. Walden, F. von, Liu, C., Aurigemma, N. & Nader, G. A. mTOR signaling regulates myotube hypertrophy by modulating protein synthesis, rDNA transcription, and chromatin remodeling. *Am J Physiol, Cell Physiol* **311,** C663–C672 (2016).

112. Abbott, J. A., Francklyn, C. S. & Robey-Bond, S. M. Transfer RNA and human disease. *Front Genet* **5,** 158 (2014).

113. Kirchner, S. & Ignatova, Z. Emerging roles of tRNA in adaptive translation, signalling dynamics and disease. *Nat Rev Genet* **16,** 98–112 (2015).

114. Grandori, C. *et al.* c-Myc binds to human ribosomal DNA and stimulates transcription of rRNA genes by RNA polymerase I. *Nat Cell Biol* **7,** 311–318 (2005).

115. Birch, J. *et al.* The initiator methionine tRNA drives cell migration and invasion leading to increased metastatic potential in melanoma. *Biology open* **5,** 1371–1379 (2016).

116. Ferguson, A. *et al.* Functional Dynamics within the Human Ribosome Regulate the Rate of Active Protein Synthesis. *Mol Cell* **60,** 475–486 (2015).

117. Subramaniam, A. R., Zid, B. M. & O'Shea, E. K. An integrated approach reveals regulatory controls on bacterial translation elongation. *Cell* **159,** 1200–1211 (2014).

118. Burnett, B. J. *et al.* Elongation factor Ts directly facilitates the formation and disassembly of the Escherichia coli elongation factor Tu·GTP·aminoacyl-tRNA

ternary complex. *J Biol Chem* **288,** 13917–13928 (2013).

119. Burnett, B. J. *et al.* Direct evidence of an elongation factor-Tu/Ts·GTP·Aminoacyl-tRNA quaternary complex. *J Biol Chem* **289,** 23917–23927 (2014).

120. Mingot, J. M., Vega, S., Cano, A., Portillo, F. & Nieto, M. A. eEF1A mediates the nuclear export of SNAG-containing proteins via the Exportin5-aminoacyl-tRNA complex. *Cell Rep* **5,** 727–737 (2013).

121. Dass, R. A. *et al.* Wnt5a Signals through DVL1 to Repress Ribosomal DNA Transcription by RNA Polymerase I. *PLoS Genet* **12,** e1006217 (2016).

122. Sarshad, A. A. *et al.* Glycogen synthase kinase (GSK) 3β phosphorylates and protects nuclear myosin 1c from proteasome-mediated degradation to activate rDNA transcription in early G1 cells. *PLoS Genet* **10,** e1004390 (2014).

123. Li, H., Tsang, C. K., Watkins, M., Bertram, P. G. & Zheng, X. F. S. Nutrient regulates Tor1 nuclear localization and association with rDNA promoter. *Nature* **442,** 1058–1061 (2006).

124. Chaillou, T., Kirby, T. J. & McCarthy, J. J. Ribosome biogenesis: emerging evidence for a central role in the regulation of skeletal muscle mass. *J Cell Physiol* **229,** 1584–1594 (2014).

125. Powers, T. & Walter, P. Regulation of ribosome biogenesis by the rapamycin-sensitive TOR-signaling pathway in Saccharomyces cerevisiae. *Mol Biol Cell* **10,** 987–1000 (1999).

126. Derenzini, M., Montanaro, L. & Trerè, D. Ribosome biogenesis and cancer. *Acta Histochem* **119,** 190–197 (2017).

127. Deisenroth, C. & Zhang, Y. Ribosome biogenesis surveillance: probing the ribosomal protein-Mdm2-p53 pathway. *Oncogene* **29,** 4253–4260 (2010).

128. McStay, B. Nucleolar organizer regions: genomic "dark matter" requiring illumination. *Genes Dev* **30,** 1598–1610 (2016).

129. McStay, B. Nucleolar dominance: a model for rRNA gene silencing. *Genes Dev* **20,** 1207–1214 (2006).